Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY
MNRAS **456,** 2662–2680 (2016) doi:10.1093/mnras/stv2833

Performance of internal covariance estimators for cosmic shear correlation functions

O. Friedrich,[1,2]★ S. Seitz,[1,2] T. F. Eifler[3,4] and D. Gruen[1,2]

[1]*University Observatory Munich, Scheinerstrasse 1, D-81679 Munich, Germany*
[2]*Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse, D-85748 Garching, Germany*
[3]*Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr, Pasadena, CA 91109, USA*
[4]*California Institute of Technology, Pasadena, CA 91125, USA*

Accepted 2015 November 30. Received 2015 November 16; in original form 2015 August 4

ABSTRACT

Data re-sampling methods such as delete-one jackknife, bootstrap or the sub-sample covariance are common tools for estimating the covariance of large-scale structure probes. We investigate different implementations of these methods in the context of cosmic shear two-point statistics. Using lognormal simulations of the convergence field and the corresponding shear field we generate mock catalogues of a known and realistic covariance. For a survey of $\sim$5000 deg^2 we find that jackknife, if implemented by deleting sub-volumes of *galaxies*, provides the most reliable covariance estimates. Bootstrap, in the common implementation of drawing sub-volumes of galaxies, strongly overestimates the statistical uncertainties. In a forecast for the complete 5-yr Dark Energy Survey, we show that internally estimated covariance matrices can provide a large fraction of the true uncertainties on cosmological parameters in a 2D cosmic shear analysis. The volume inside contours of constant likelihood in the $\Omega_{\rm m}$–σ_8 plane as measured with internally estimated covariance matrices is on average $\gtrsim$85 per cent of the volume derived from the true covariance matrix. The uncertainty on the parameter combination $\Sigma_8 \sim \sigma_8\Omega_{\rm m}^{0.5}$ derived from internally estimated covariances is $\sim$90 per cent of the true uncertainty.

Key words: methods: data analysis – methods: statistical – cosmological parameters – large-scale structure of Universe.

1 INTRODUCTION

Two-point statistics of cosmological random fields such as the cosmic shear correlation functions or the galaxy clustering angular correlation function are common probes of the large-scale structure of the Universe. Recent measurements of these correlation functions are e.g. reported in Thomas, Abdalla & Lahav (2011), Kilbinger et al. (2013), de Simoni et al. (2013) and Becker et al. (2015). In order to use these statistics for constraining cosmological models one needs a quantitative description of the joint distribution of the correlation function estimators. When assuming multivariate Gaussian errors, this is given by the covariance matrix. On large angular scales, this covariance matrix can – both for cosmic shear and galaxy clustering – be well described by a Gaussian approximation for the involved fields (Schneider et al. 2002; Crocce, Cabré & Gaztañaga 2011). It has, however, been shown that the Gaussian approximation fails to describe the true probability distribution function (PDF) of the weak lensing convergence field (Taruya et al. 2002; Vale & White 2003) and that it underestimates the true covariance of the cosmic shear correlation functions on small scales, which can be alleviated by an empirical re-scaling (Semboloni et al. 2007; Sato et al. 2011), a lognormal approximation (Hilbert, Hartlap & Schneider 2011), or halo model approaches (e.g. Cooray & Hu 2001; Takada & Jain 2009; Eifler et al. 2014).

Alternatives to modelling the covariance matrix are to estimate it from many independent realizations of cosmological *N*-body simulations or to estimate it internally, i.e. from the data itself. The latter method is independent of assuming a particular cosmological model and is hence often used to complement the other methods (Kilbinger et al. 2013; Wang, Brunner & Dolence 2013; Becker et al. 2015.).

So far the performance of internal covariance estimators has only been systematically studied for the galaxy clustering two-point function (in most detail by Norberg et al. 2009) or for cross-correlations of the cosmic microwave background (CMB) and the galaxy field (Cabré et al. 2007). In our paper, we will concentrate on cosmic shear correlation functions. We will show that the shape noise part of the covariance can be very accurately estimated internally, while the cosmic variance part is generally underestimated. Gaussian simulations of the convergence field would hence yield an overly optimistic test of internal covariance estimators, since the

★E-mail: oliverf@usm.uni-muenchen.de

Gaussian model underpredicts the cosmic variance contribution to the covariance. We overcome this problem by employing lognormal simulations of the convergence field.

In our paper, we want to study the performance of internal covariance estimators such as bootstrap, jackknife or the sub-sample covariance. There is no complete agreement in the literature yet on whether internal covariance estimates can be used to constrain cosmological parameters from measured two-point correlations or whether they are a mere tool to generate reasonable error bars in plots of correlation functions (see e.g. Norberg et al. 2009; de Simoni et al. 2013; Taylor, Joachimi & Kitching 2013; Wang et al. 2013). We want to address the questions of how many internal resamplings are required in order to get a stable covariance matrix, whether internal estimators over- or underestimate the covariance matrix and whether/how internal covariance estimates can yield unbiased estimates of the *inverse* covariance matrix.

Our paper is organized as follows: In Section 2, we introduce the cosmic shear correlation functions and explain the Gaussian and the lognormal model for the covariance of two-point function estimators. In Section 3, we describe the simulations we use to generate mock shape catalogues that follow any given input power spectrum and whose underlying convergence field has a lognormal PDF. These are the simulations with which we will test the performance of internal covariance estimators.

In Section 4, we introduce two distinct ways of performing internal estimation of the covariance of two-point measures – the pair scheme and the galaxy scheme. Furthermore, we are explaining why jackknife, bootstrap and sub-sample covariance are almost equivalent.

In Section 5, we apply internal covariance estimators to simulated cosmic shear surveys. We show that in the pair scheme all estimators are almost identical and we demonstrate the systematic effects of the different estimation schemes when varying the number of resamplings. Our method to find optimal estimation schemes has to be re-run for any specific survey, because the performance of internal estimators depends crucially on the depth and area of a survey. In the end of Section 5 we configure our simulations to match the complete, 5-yr Dark Energy Survey (DES; Flaugher 2005; The Dark Energy Survey Collaboration 2005) and test the accuracy of jackknife covariance matrices for this particular setting. The code used for our simulations is made publicly available.[1]

In Section 6, we discuss the results of our work.

2 COSMIC SHEAR BASICS

2.1 Cosmic shear correlation functions

Cosmic shear measures the correlated distortion of galaxy shapes due to gravitational lensing by the large-scale structure of the universe as a function of the angular distance of galaxy pairs on the sky. We follow here the notation of Schneider et al. (2002) and employ the flat-sky-approximation, i.e. we assume a tangential Cartesian coordinate system $\boldsymbol{\vartheta} = (\vartheta_1, \vartheta_2)$ on the sky.

In this coordinate system the comic shear field is at each point characterized by a complex number $\boldsymbol{\gamma}(\boldsymbol{\vartheta}) = \gamma_1 + \mathrm{i}\gamma_2$. If the separation vector $\Delta\boldsymbol{\vartheta} = \boldsymbol{\vartheta}_2 - \boldsymbol{\vartheta}_1$ of two points on the sky has the polar angle ϕ then the *tangential* and *cross*-components of $\boldsymbol{\gamma}$ at $\boldsymbol{\vartheta}_2$ and $\boldsymbol{\vartheta}_1$ (with respect to each other) are defined as

$$\gamma_t = -\mathcal{R}\mathrm{e}\left(\boldsymbol{\gamma}\mathrm{e}^{-2\mathrm{i}\phi}\right); \qquad \gamma_\times = -\mathcal{I}\mathrm{m}\left(\boldsymbol{\gamma}\mathrm{e}^{-2\mathrm{i}\phi}\right). \tag{1}$$

The *cosmic shear correlation functions* $\xi_\pm(\theta)$ are defined as the expectation values

$$\xi_\pm(\theta) = \langle\gamma_{t,1}\gamma_{t,2}\rangle \pm \langle\gamma_{\times,1}\gamma_{\times,2}\rangle, \tag{2}$$

where θ is the absolute value of $\Delta\boldsymbol{\vartheta}$. It can be computed in terms of the power spectrum $P_\kappa(\ell)$ of the scalar *convergence field* $\kappa(\boldsymbol{\vartheta})$ as

$$\xi_\pm(\theta) = \frac{1}{2\pi}\int \mathrm{d}\ell\ \ell\ P_\kappa(\ell) J_{0/4}(\ell\theta), \tag{3}$$

where $J_0(x)$ ($J_4(x)$) is the zeroth-order (fourth-order) Bessel function.

The shape of a galaxy can be characterized by a complex number $\boldsymbol{\epsilon}$ which is to first order the sum of the intrinsic shape $\boldsymbol{\epsilon}^{\mathrm{in}}$ of the galaxy and the distortion caused by gravitational lensing, i.e. the value $\boldsymbol{\gamma}(\boldsymbol{\vartheta})$ at the location $\boldsymbol{\vartheta}$ of the galaxy,

$$\boldsymbol{\epsilon} = \boldsymbol{\epsilon}^{\mathrm{in}} + \boldsymbol{\gamma}. \tag{4}$$

In a cosmic shear survey, the shapes $\boldsymbol{\epsilon}_i$ of many galaxies are measured and (cf. Schneider et al. 2002) an estimator for the correlation functions can be constructed as

$$\hat{\xi}_\pm(\theta) = \frac{\sum_{ij} w_j w_j (\epsilon_{t,i}\epsilon_{t,j} \pm \epsilon_{\times,i}\epsilon_{\times,j})\Delta_\theta(ij)}{\sum_{ij} w_j w_j \Delta_\theta(ij)}, \tag{5}$$

where we have allowed for some weighting scheme w_i for the shape measurements and where the filter $\Delta_\theta(ij)$ selects all galaxy pairs (i, j) in the survey whose angular separation lies in some finite bin around θ. The normalization in equation (5) is the effective number of galaxy pairs in a bin around θ, which we will abbreviate as

$$N_{\mathrm{p}}(\theta) = \sum_{ij} w_j w_j \Delta_\theta(ij). \tag{6}$$

2.2 Covariance of the correlation functions

The covariance matrix of the estimator in equation (5) is defined as

$$\begin{aligned} C_{\pm,\pm}(\theta_1,\theta_2) &= \langle(\hat{\xi}_\pm(\theta_1) - \xi_\pm(\theta_1))(\hat{\xi}_\pm(\theta_2) - \xi_\pm(\theta_2))\rangle \\ &= \langle\hat{\xi}_\pm(\theta_1)\hat{\xi}_\pm(\theta_2)\rangle - \xi_\pm(\theta_1)\xi_\pm(\theta_2). \end{aligned} \tag{7}$$

In order to compute this covariance matrix, it is convenient to split $\xi_\pm(\theta)$ into the three different contribution

$$\begin{aligned} \hat{\xi}^{\mathrm{nn}}_\pm(\theta) &= \frac{\sum_{ij} w_i w_j (\epsilon^{\mathrm{in}}_{t,i}\epsilon^{\mathrm{in}}_{t,j} \pm \epsilon^{\mathrm{in}}_{\times,i}\epsilon^{\mathrm{in}}_{\times,j})\Delta_\theta(ij)}{N_{\mathrm{p}}(\theta)}, \\ \hat{\xi}^{\mathrm{ss}}_\pm(\theta) &= \frac{\sum_{ij} w_i w_j (\gamma_{t,i}\gamma_{t,j} \pm \gamma_{\times,i}\gamma_{\times,j})\Delta_\theta(ij)}{N_{\mathrm{p}}(\theta)}, \\ \hat{\xi}^{\mathrm{sn}}_\pm(\theta) &= \frac{\sum_{ij} w_i w_j (\epsilon^{\mathrm{in}}_{t,i}\gamma_{t,j} \pm \epsilon^{\mathrm{in}}_{\times,i}\gamma_{\times,j})\Delta_\theta(ij)}{N_{\mathrm{p}}(\theta)} \end{aligned} \tag{8}$$

which are the autocorrelation of the intrinsic shape noise, the autocorrelation of the shear signal and their cross-correlation. The whole estimator (5) is given in terms of these as

$$\hat{\xi}_\pm(\theta) = \hat{\xi}^{\mathrm{nn}}_\pm(\theta) + \hat{\xi}^{\mathrm{ss}}_\pm(\theta) + 2\cdot\hat{\xi}^{\mathrm{sn}}_\pm(\theta). \tag{9}$$

Under the assumption that the shear signal and the shape noise are independent of each other it is obvious that

$$\langle\hat{\xi}^{\mathrm{nn}}_\pm(\theta_1)\hat{\xi}^{\mathrm{sn}}_\pm(\theta_2)\rangle = 0 = \langle\hat{\xi}^{\mathrm{ss}}_\pm(\theta_1)\hat{\xi}^{\mathrm{sn}}_\pm(\theta_2)\rangle. \tag{10}$$

[1] www.usm.uni-muenchen.de/people/oliverf/, the code also contains many other useful features, that e.g. enable the user to create mock data suitable for galaxy–galaxy lensing or galaxy clustering measurements.

If the intrinsic shape of any two galaxies is assumed to be uncorrelated, we can also conclude that

$$\langle\hat{\xi}_{\pm}^{\rm nn}\rangle = 0 \quad \text{for} \quad \theta > 0 \tag{11}$$

and hence

$$\langle\hat{\xi}_{\pm}^{\rm nn}(\theta_1)\hat{\xi}_{\pm}^{\rm ss}(\theta_2)\rangle = \langle\hat{\xi}_{\pm}^{\rm nn}(\theta_1)\rangle \cdot \langle\hat{\xi}_{\pm}^{\rm ss}(\theta_2)\rangle = 0 \quad \text{for} \quad \theta_1, \theta_2 > 0.$$

The covariance matrix can thus be split into three different contributions,

$$\mathbf{C}_{\pm,\pm} = \mathbf{C}_{\pm,\pm}^{\rm nn} + \mathbf{C}_{\pm,\pm}^{\rm ss} + \mathbf{C}_{\pm,\pm}^{\rm sn}, \tag{12}$$

namely

$$\begin{aligned} C_{\pm,\pm}^{\rm nn}(\theta_1,\theta_2) &= \langle\hat{\xi}_{\pm}^{\rm nn}(\theta_1)\hat{\xi}_{\pm}^{\rm nn}(\theta_2)\rangle, \\ C_{\pm,\pm}^{\rm ss}(\theta_1,\theta_2) &= \langle\hat{\xi}_{\pm}^{\rm ss}(\theta_1)\hat{\xi}_{\pm}^{\rm ss}(\theta_2)\rangle - \xi_{\pm}(\theta_1)\xi_{\pm}(\theta_2), \\ C_{\pm,\pm}^{\rm sn}(\theta_1,\theta_2) &= 4\cdot\langle\hat{\xi}_{\pm}^{\rm sn}(\theta_1)\hat{\xi}_{\pm}^{\rm sn}(\theta_2)\rangle. \end{aligned} \tag{13}$$

The $\mathbf{C}_{\pm,\pm}^{\rm ss}$ term depends on four-point functions of the shear field and is called the *cosmic variance* term. In order to evaluate it, further assumptions on the PDF of the shear or the convergence field are needed and we will discuss two possible models for the convergence PDF in Sections 2.2.1 and 2.2.2 – the Gaussian and the lognormal model.

The contributions $\mathbf{C}_{\pm,\pm}^{\rm nn}$ and $\mathbf{C}_{\pm,\pm}^{\rm sn}$ can be computed without additional assumptions. In Joachimi, Schneider & Eifler (2008) it is derived that they are given by[2]

$$\begin{aligned} C_{\pm\pm}^{\rm sn}(\theta_1,\theta_2) &= \frac{\sigma_\epsilon^2}{\pi A\bar{n}}\int {\rm d}\ell\ \ell\ J_{0/4}(\ell\theta_1)\ J_{0/4}(\ell\theta_2)\ P_\kappa(\ell), \\ C_{++}^{\rm nn}(\theta_1,\theta_2) &= C_{-}^{\rm nn}(\theta_1,\theta_2) \\ &= \frac{\sigma_\epsilon^4}{N_{\rm p}(\theta_1)}\cdot\delta_{\theta_1,\theta_2}, \\ C_{+-}^{\rm nn}(\theta_1,\theta_2) &= 0, \end{aligned} \tag{14}$$

where A is the survey area, $\bar{n}$ is the number density of galaxies, σ_ϵ is the dispersion of the intrinsic ellipticity which is defined as

$$\sigma_\epsilon^2 := \langle\boldsymbol{\epsilon}^{\rm in}\boldsymbol{\epsilon}^{\rm in*}\rangle, \tag{15}$$

and P_κ is again the convergence power spectrum.

2.2.1 Gaussian approximation

In the paper series by Schneider et al. (2002), Kilbinger & Schneider (2004) and Joachimi et al. (2008), the covariance matrix is studied in the Gaussian approximation, i.e. assuming that the convergence field has a Gaussian PDF such that its four-point correlation functions can be expressed in terms of its two-point correlation functions.

For the case where the survey geometry is much larger than the angular scales considered in the correlation functions, Joachimi et al. (2008) derive the following expressions for the cosmic variance term:

$$C_{\pm\pm}^{\rm ss}(\theta_1,\theta_2) = \frac{1}{\pi\,A}\int {\rm d}\ell\ \ell\ J_{0/4}(\ell\theta_1)\ J_{0/4}(\ell\theta_2)\ P_\kappa^2(\ell). \tag{16}$$

However, due to the finite geometry of any given survey equation (16) generally overestimates the covariance of Gaussian field as was demonstrated in Sato et al. (2011). This *finite area effect* according to Sato et al. is not important for surveys larger than 1000 deg^2. For smaller surveys a method developed in Kilbinger & Schneider (2004) which does not employ an ensemble average over galaxy positions should be used to evaluate the Gaussian covariance. This method was for example used in the analysis of CHFTLenS data in Kilbinger et al. (2013). The finite area effect is also important for internal covariance estimation and will be further discussed in Section 4.2.

2.2.2 Shifted lognormal approximation

As e.g. reported by Taruya et al. (2002), Vale & White (2003) or by Hilbert et al. (2011), the Gaussian model fails to describe the true PDF of the convergence and especially on small separations poorly represents the true covariance of the cosmic shear two-point functions.

Hilbert et al. (2011) propose a different model for the convergence PDF, namely that of a *zero-mean shifted lognormal distribution*. In this approach, the convergence at a given point on the sky is assumed to be of the form

$$\kappa(\boldsymbol{\theta}) = \exp[n(\boldsymbol{\theta})] - \kappa_0, \tag{17}$$

where $n(\boldsymbol{\theta})$ is a Gaussian random field (not necessarily with a vanishing mean) and the *minimal convergence parameter* κ_0 is chosen such that $\langle\kappa\rangle = 0$. Hilbert et al. (2011) show that from the corresponding PDF a model for the shear–shear contribution to the covariance matrix can be derived. Considering only the most important terms they also provide a simplified lognormal covariance, which reads

$$\begin{aligned} C_{\pm\pm}^{\rm ss}(\theta_1,\theta_2) = \frac{1}{\pi A}\int {\rm d}\ell\ \ell\ J_{0/4}(\ell\theta_1)\ J_{0/4}(\ell\theta_2)\ P_\kappa^2(\ell) \\ + \frac{8\pi}{\kappa_0^2 A}\,\xi_{\pm}(\theta_1)\xi_{\pm}(\theta_2)\int_0^{\theta_{\rm A}} {\rm d}\theta\ \theta\ \xi_{+}(\theta), \end{aligned} \tag{18}$$

where θ_A represents the 'radius' of the survey, given by

$$\theta_{\rm A} = \sqrt{\frac{A}{\pi}}. \tag{19}$$

Comparing equation (18) to equation (16) one can see that the simplified lognormal approximation to $\mathbf{C}_{\pm\pm}^{\rm ss}$ consists of only one correction term to the Gaussian model. In our paper, we will simulate lognormally distributed convergence fields and use equation (18) to compute the cosmic variance part of our model covariance.

2.2.3 Finite bin width

The expressions presented above for the covariance of $\hat{\xi}_{\pm}$ are derived under the assumption of small angular bins (Schneider et al. 2002). However, in Section 5.2 we need correct covariance expressions also for data vectors where the relative bin width is $\sim$0.3, i.e. where the assumption of small bins does not hold. This is in fact the more realistic case, since broad bins are commonly used to reduce the number of data points (see e.g. Kilbinger et al. 2013; Becker et al. 2015).

Hence, in Section 5.2 we proceed as follows: we first compute the lognormal model for the covariance, equation (18), for a set of very small angular bins $\tilde{\theta}_i$, $i = 1, \ldots, \tilde{N}$. Then we apply a linear transformation that takes the large data vector of the small angular bins to a smaller data vector by putting together p neighbouring bins

[2] As in Schneider et al. (2002) they employ an ensemble average over the galaxy positions to derive their expressions.

of the old data vector,

$$\theta_j = \sum_{i=p\cdot(j-1)+1}^{p\cdot j} \tilde{\theta}_i N_\mathrm{p}(\tilde{\theta}_i) \Bigg/ \sum_{i=p\cdot(j-1)+1}^{p\cdot j} N_\mathrm{p}(\tilde{\theta}_i)$$

$$\hat{\xi}(\theta_j) = \sum_{i=p\cdot(j-1)+1}^{p\cdot j} \hat{\xi}(\tilde{\theta}_i) N_\mathrm{p}(\tilde{\theta}_i) \Bigg/ \sum_{i=p\cdot(j-1)+1}^{p\cdot j} N_\mathrm{p}(\tilde{\theta}_i), \quad (20)$$

where $N_\mathrm{p}(\tilde{\theta}_i)$ is the number of pairs in the ith bin of the finer data vector.

The same linear transformation is then applied to the covariance matrix of the large data vector to get the covariance matrix of the compressed data vector. Hence, if the transformation (20) is represented by the $(N/p) \times N$ matrix $\mathbf{A}$, the covariance matrix of the compressed data vector is given by

$$\mathbf{C}_\mathrm{comp} = \mathbf{A}\,\mathbf{C}\,\mathbf{A}^\mathrm{T}, \quad (21)$$

where $\mathbf{A}^\mathrm{T}$ is the transposed transformation. We find that for $\hat{\xi}_-$ this decreases the mixed and cosmic variance part of the covariance by $\gtrsim$30 per cent, while for $\hat{\xi}_+$ it makes almost no difference. The reason is that adjacent bins in ξ_+ are much more correlated than adjacent bins in ξ_-. Hence, if two bins in ξ_+ are joined, the variance of the joined bin is almost identical to that of the individual bins and equation (18) can still be applied.[3]

3 LOGNORMAL SIMULATIONS

Simon, King & Schneider (2004) describe a quick method to simulate cosmic shear surveys based on a Gaussian convergence field for any given convergence power spectrum. On a quadratic grid in 2D-Fourier space they generate at each point $\boldsymbol{\ell}$ of the grid a value of the convergence

$$\hat{\kappa}(\boldsymbol{\ell}) = \kappa_1(\boldsymbol{\ell}) + i\kappa_2(\boldsymbol{\ell}), \quad (22)$$

where the components $\kappa_i(\boldsymbol{\ell})$ are drawn from a Gaussian distribution with zero mean and variance

$$\sigma_\ell^2 = \frac{1}{2V} P_\kappa(\ell). \quad (23)$$

Here P_κ is the desired convergence power spectrum and V is the volume of the grid in angular space which is given in terms of the grid spacing $\Delta\ell$ as

$$V = \left(\frac{2\pi}{\Delta\ell}\right)^2. \quad (24)$$

In order to achieve a convergence field that is real valued in angular space one has to impose the condition

$$\hat{\kappa}(\boldsymbol{\ell}) = \hat{\kappa}^*(-\boldsymbol{\ell}).$$

In Fourier space the shear field is related to the convergence field by the equation[4]

$$\hat{\gamma}(\boldsymbol{\ell}) = \frac{\ell_1^2 - \ell_2^2 + 2i\ell_1\ell_2}{\ell^2}\hat{\kappa}(\boldsymbol{\ell}). \quad (25)$$

A Fourier transform then gives the shear field in angular space.

The main idea in generating a lognormal random field is to generate a Gaussian field $n(\boldsymbol{\theta})$ with the method of Simon et al. (2004) and transform it into $\kappa(\boldsymbol{\theta})$ via equation (17). According to Martin, Schneider & Simon (2012) and Takahashi et al. (2014) the power spectrum of $n(\boldsymbol{\theta})$, P_n, can be computed from P_κ as follows.

First, the two-point function of $\kappa(\boldsymbol{\theta})$ is given in terms of the power spectrum P_κ by

$$\xi_\kappa(\theta) = \frac{1}{2\pi}\int_0^\infty \mathrm{d}\ell\ \ell\ P_\kappa(\ell)\ J_0(\ell\theta). \quad (26)$$

Next, the two-point function ξ_κ is related to the two-point function of $n(\boldsymbol{\theta})$ via (see e.g. equation B.8 of Hilbert et al. 2011)

$$\xi_n(\theta) = \ln\left(\xi_\kappa(\theta)/\kappa_0^2 + 1\right), \quad (27)$$

where κ_0 is the minimal convergence parameter from equation (17). Finally, the power spectrum of the Gaussian field $n(\boldsymbol{\theta})$ is found by

$$P_n(\ell) = 2\pi\int_0^\infty \mathrm{d}\theta\ \theta\ \xi_n(\theta)\ J_0(\ell\theta). \quad (28)$$

The field $n(\boldsymbol{\theta})$ can now be generated as described by Simon et al. (2004). However, this way $n(\boldsymbol{\theta})$ will have a mean value of zero. In order to ensure that $\langle\kappa\rangle = 0$ the mean value

$$\mu = \kappa_0 - \frac{\sigma^2}{2} \quad (29)$$

has to be added, where σ^2 is the variance of the Gaussian field. The convergence field $\kappa(\boldsymbol{\theta})$ now has to be transformed into Fourier space. Using equation (25) one can then compute the Fourier modes of the shear field and another Fourier transform gives the desired shear field in angular space.

3.1 Setup and validation of the simulations

The harmonic space grid we are using has a total number of $(2^{16})^2$ grid points and a grid spacing of $\Delta\ell = 2$. Hence, in each axis it ranges from $-\ell_\mathrm{max} = -2^{16}$ to $+\ell_\mathrm{max} = +2^{16}$. All modes $\boldsymbol{\gamma}(\boldsymbol{\ell})$ with $|\boldsymbol{\ell}| > \ell_\mathrm{max}$ (i.e. the corners of the grid) are set to zero. The mode $\boldsymbol{\gamma}(0)$ is also set to 0 and all other modes are generated as explained above. Especially, we have to fix a cosmology and assume a certain redshift distribution of sources, $p(z)$, to compute the convergence power spectrum P_κ.

Following equation (24) the grid in angular space has a volume of $V = 2\pi/2 \approx 10^4\,\mathrm{deg}^2$. We cut a sub-grid of size A out of the centre of that volume. Onto this sub-grid, we uniformly place galaxies with a certain number density n_gal. The shear of each individual galaxy is then determined by quadratic interpolation of the grid onto the galaxy position. Finally, a Gaussian intrinsic shape noise with an ellipticity dispersion σ_ϵ is added to get the total shape of the galaxy. Note that we simply add the shear signal and intrinsic ellipticity, hereby ignoring the effects of *reduced shear*.

In this work, we always keep the cosmology fixed to that used by Hilbert et al. (2011), i.e. a flat ΛCDM universe with $(\Omega_\mathrm{m}, \Omega_\mathrm{b}, \sigma_8, h_{100}, n_\mathrm{s}) = (0.25, 0.045, 0.9, 0.73, 1.0)$. To compute the convergence power spectrum, we employ *halofit* (Smith et al. 2003) using the open source code NICAEA.[5] The source distribution is taken to have the form

$$p(z) = \frac{3z^2}{2z_0^3}\,\mathrm{e}^{-\left(\frac{z}{z_0}\right)^{3/2}}, \quad \text{where } z_0 = \frac{z_\mathrm{median}}{1.412}. \quad (30)$$

This is the same form that was also assumed by Hilbert et al. (2011). The ellipticity dispersion is always set to 0.3 per component, i.e.

[3] A similar reasoning can be applied for the off-diagonal terms of the covariance.

[4] See equation 2.1.11 of Kaiser & Squires (1993) or equation 25 of Simon et al. (2004).

[5] By Kilbinger et al., see also Kilbinger et al. (2009) or www.cosmostat.org/software/nicaea/

Table 1. The different configurations of mock catalogues used in this paper. The last column also gives the number of realizations that were generated for our analyses. Setup IIa was simulated 1000 times in order to get a reliable reference covariance matrix. The other setups are large enough in area to trust the approximations made in the lognormal model covariance.

Setup	A (deg)	$n_{\rm gal}$(arcmin^{-2})	$z_{\rm median}$	κ_0	$N_{\rm sim}$
I	4900	20	1.0	0.032	50
IIa	∼150	6	0.7	0.019	1000
IIb	5000	6	0.7	0.019	10
IIc	5000	10	0.7	0.019	10

$\sigma_\epsilon = \sqrt{2} \times 0.3$. All other quantities, i.e. area A, source density $n_{\rm gal}$ and median redshift $z_{\rm median}$, will be varied throughout Section 5. The different setups are summarized in Table 1.

The redshift distribution of setup I is exactly that of Hilbert et al. (2011) and imitates a rather deep survey comparable e.g. to *Euclid* (Laureijs et al. 2011). In this setup, we measure the two-point correlation functions in 35 logarithmic bins from $\theta_{\rm min} = 1$ arcmin to $\theta_{\rm max} = 150$ arcmin. The correlation function measurement is in all setups carried out using the TREECORR tree code.[6] The area A is taken to be a square of 70 deg ×70 deg. The minimal convergence parameter κ_0 is chosen to be 0.032 as suggested by Hilbert et al. for this redshift distribution.

The area, galaxy density and redshift distribution of setup IIa are chosen to be similar to that of DES science verification data (DES-SV) which was used in Becker et al. (2015). In this setup, we measure the two-point correlation functions in 15 logarithmic bins from $\theta_{\rm min} = 2$ arcmin to $\theta_{\rm max} = 300$ arcmin, which is the data vector used by Becker et al. (2015). We also reproduce the irregular shape of DES-SV, i.e. we use an SV-shaped HEALPIX mask to cut out the sub-volume A.

The setups IIb and IIc are aimed at a forecast for the final 5-yr DES data. In IIb, we are assuming the same source density as in DES-SV and in IIc a slightly higher one. Note that in principle, when adjusting the source density, we should also adjust the median redshift of the sources. But we will ignore this point, since our redshift distribution is anyway only a rough match to that of DES. Thus, for all the setups IIa, IIb and IIc we take a median redshift of 0.7. Furthermore, for all these setups we use the empirical relation $\kappa_0(z)$ found by Hilbert et al. (2011) to fix the minimal convergence parameter. Inserting the mean redshift of $z_{\rm mean} \approx 0.745$ gives a value of $\kappa_0 = 0.019$. The area in setups IIb and IIc are simply taken to be square shaped.

To validate our simulations, we generate 1000 independent realizations of setup I. In order to speed up the computations, we decrease the number of galaxies with respect to our jackknife analysis by a factor of 5, i.e. to $n_{\rm gal} = 4$ arcmin^{-2}, while at the same time decreasing the ellipticity dispersion by a factor of $\sqrt{5}$. This way the covariance expressions in equation (18) are unaffected.

In Fig. 1, we show the mean measured correlation functions in the mock surveys. The measured two-point functions and those derived from the input model agree well on most scales. Only at small angular scales, the measured value of ξ_- differs significantly from the input model. The reason is the artificial cut-off at high ℓ-values which introduces artefacts both in the model and the simulation – as can be seen from the oscillatory behaviour of ξ_-. To keep our analyses in Section 5 unaffected by these artefacts, we will only consider those bins in ξ_- that have $\theta \gtrsim 4.5$ arcmin. For ξ_+ we continue to use a range of 1 arcmin $< \theta <$ 150 arcmin. Also, for the setups IIa to IIc (not shown here) the discrepancy in ξ_- turns out to be less significant. Hence for these setups we stay with $\theta_{\rm min} = 2$ arcmin.

[6] see Jarvis et al. (2004) or github.com/rmjarvis/TreeCorr

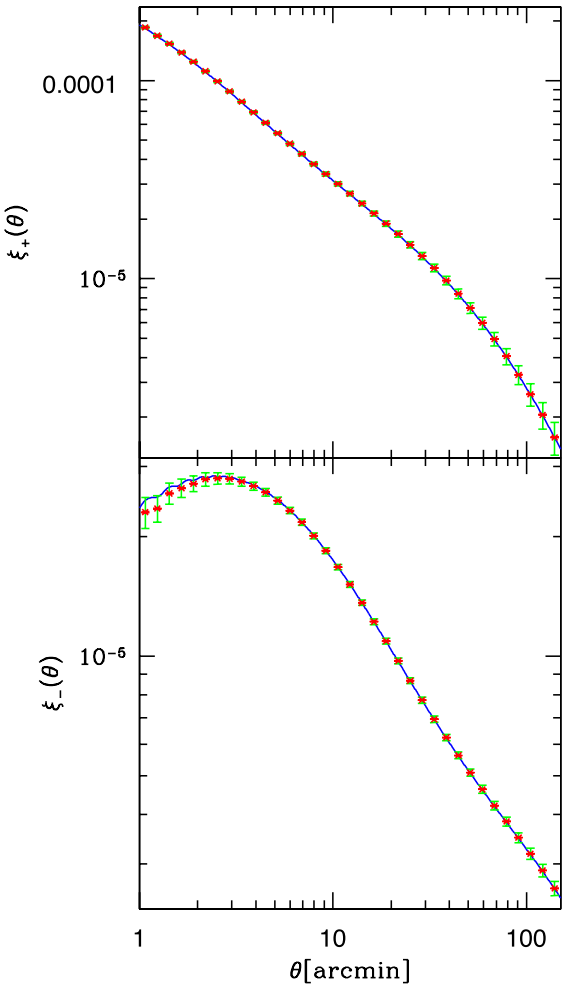

Figure 1. Comparison of the mean correlation functions from 1000 simulations (red dots) and the input model (blue line). The red error bars show the standard deviation of the mean and the green error bars show the standard deviation of the single correlation function measurements. We used the redshift distribution of Hilbert et al. (2011) to compute the input power spectrum and we also used their value of κ_0 to generate the lognormal convergence. Note that in Section 5.2, we will use a different configuration.

To test the higher order statistical properties of our simulations, we computed the sample covariance of the 1000 mock surveys. Fig. 2 compares this sample covariance to the predictions from equation (18). The relative deviation between measured variance and the lognormal model is ≤20 per cent for ξ_+ and ≤15 per cent for ξ_-. For both correlation functions these deviations seem to be significant given the uncertainties of the sample covariance estimate. However, the sample variance values at different angular scales are highly correlated, which makes a 'χ-by-eye' judgement of the fit impossible. When transforming the covariance matrices into the eigenbasis of the model covariance (right-hand panel of Fig. 2), the variance values become uncorrelated and the agreement of the covariance matrices becomes more evident. The eigenvalues at which the lognormal covariance significantly differs from the sample covariance of our simulations are three orders of magnitude smaller than the biggest eigenvalues for ξ_+ and more than two

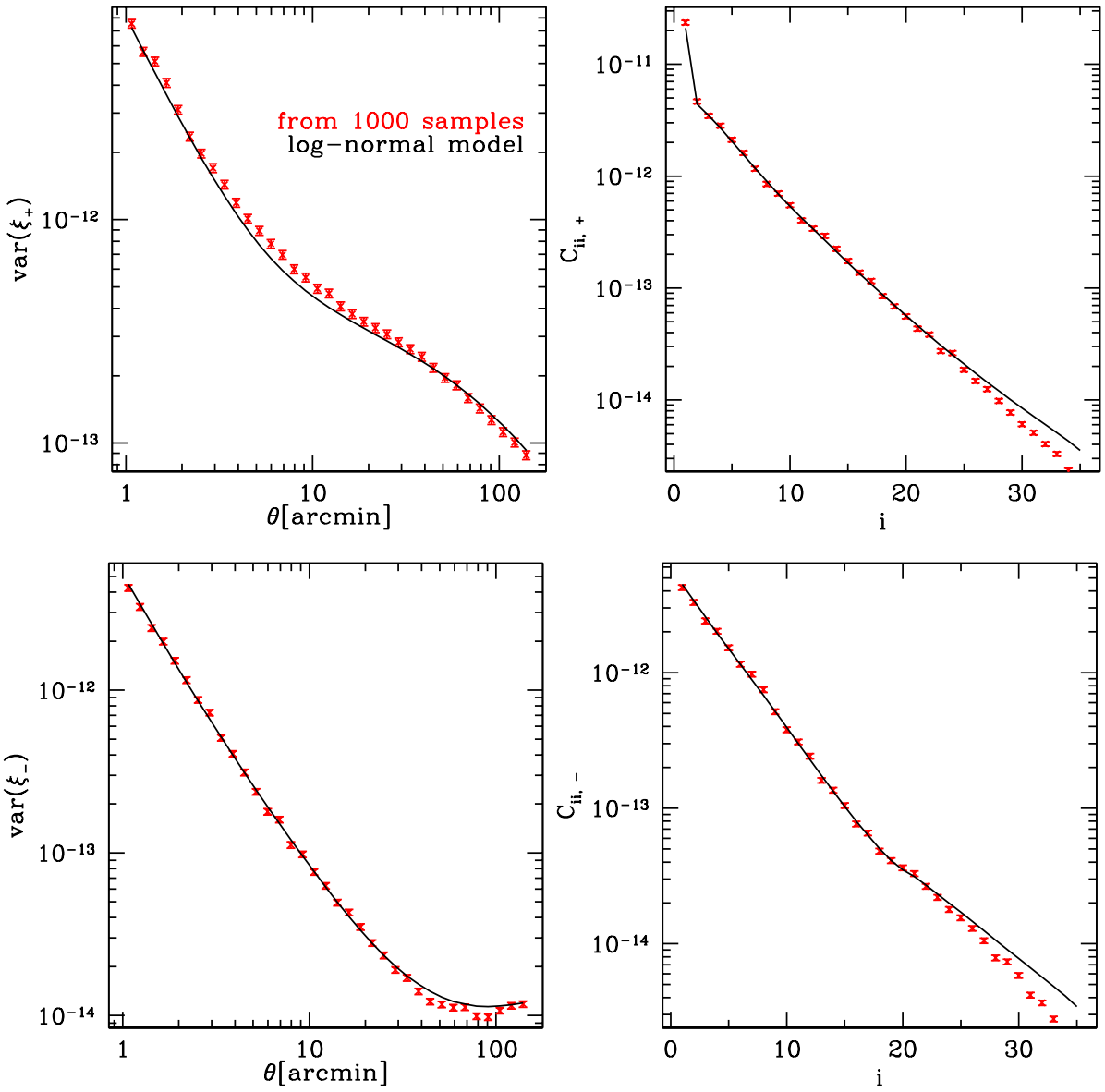

Figure 2. Left: sample variance from 1000 independent simulations compared to the lognormal input model. The error bars are assuming a Wishart distribution, note however that the different sample variance values are correlated. Right: in the diagonal basis of the model covariance matrix, the sample variance values should independently follow a χ^2-distribution. The model and the simulations are consistent for the ≈20 largest eigenvalues of the model covariance matrix.

orders of magnitude smaller than the biggest eigenvalues for ξ_- (cf. right-hand panel of Fig. 2). Finally, our analyses in Section 5 remain unchanged when the lognormal covariance is exchanged by the sample covariance of the 1000 independent realizations, which validates the simulations for our purposes.

4 INTERNAL COVARIANCE ESTIMATION FOR TWO-POINT CORRELATION FUNCTIONS

Suppose the correlation functions $\xi_\pm$ have been measured in finite bins around a set of angular distances $\theta_i, i = 1, \ldots, d$. Let $\hat{\boldsymbol{\xi}}$ be either one of the data vectors $[\xi_\pm(\theta_1), \ldots, \xi_\pm(\theta_d)]$ or the joint data vector of both correlation functions.

If $\boldsymbol{\xi}[\boldsymbol{\pi}]$ is a model for the measurement $\hat{\boldsymbol{\xi}}$ which depends on a set of parameters $\boldsymbol{\pi}$, then a common statistic for constraining the possible values of $\boldsymbol{\pi}$ is the χ^2 statistic (Kilbinger & Schneider 2004), i.e.

$$\chi^2[\boldsymbol{\pi}] = (\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}[\boldsymbol{\pi}])^{\mathrm{T}} \mathbf{C}^{-1} (\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}[\boldsymbol{\pi}]), \tag{31}$$

where $\mathbf{C}$ is the covariance matrix of $\hat{\boldsymbol{\xi}}$. One way to get the covariance matrix is to model it theoretically. As we have seen in Section 2.2, the modelling of the covariance depends crucially on the PDF of the convergence field (Schneider et al. 2002; Hilbert et al. 2011; Sato et al. 2011) and neither the Gaussian nor the lognormal approximation match a realistic convergence PDF. Also, the model covariance matrix will depend on cosmological parameters itself which, at least for small surveys, has to be taken into account when deriving parameter constraints (Eifler, Schneider & Hartlap 2009).

A way to get around modelling the covariance matrix directly is to use the sample covariance of measurements of the correlation functions in a set of independent N-body simulations (cf. Sato et al. 2009; Takahashi et al. 2009; Hilbert et al. 2011; Harnois-Déraps &

van Waerbeke 2015 or for an application to data Kilbinger et al. 2013) which however still depends on the model parameters, i.e. on the assumption of a particular cosmological model. Another alternative to modelling the covariance matrix is to estimate it from the data itself. In the following, we will introduce three different internal covariance estimation methods – the *sub-sample covariance*, the *delete-one-jackknife* and the *bootstrap* (cf. Loh 2008; Norberg et al. 2009).

4.1 Sub-sample covariance

Let us split the area A of our cosmic shear survey into N equally shaped and sized sub-regions of the area $A_{\rm sub} = A/N$. In each sub-region $\alpha = 1, \ldots, N$, a measurement of the data vector $\hat{\boldsymbol{\xi}}^\alpha$ can be carried out. Assuming that each sub-region has approximately the same number of galaxies and that the correlation functions are measured on scales much smaller than $\sqrt{A_{\rm sub}}$, the measurement of $\hat{\boldsymbol{\xi}}$ in the whole survey is given by

$$\hat{\boldsymbol{\xi}} \approx \bar{\boldsymbol{\xi}} := \frac{1}{N}\sum_{\alpha=1}^{N} \hat{\boldsymbol{\xi}}^\alpha, \tag{32}$$

i.e. it is the mean value of the measurements in the sub-regions. If the measurements $\hat{\boldsymbol{\xi}}^\alpha$ are independent, then the ij-th element of their covariance matrix can be estimated by

$$\langle \Delta\hat{\boldsymbol{\xi}}_i^\alpha \Delta\hat{\boldsymbol{\xi}}_j^\alpha \rangle \approx \frac{1}{N-1}\sum_{\beta=1}^{N} (\hat{\boldsymbol{\xi}}^\beta - \bar{\boldsymbol{\xi}})_i \, (\hat{\boldsymbol{\xi}}^\beta - \bar{\boldsymbol{\xi}})_j, \tag{33}$$

where $\Delta\hat{\boldsymbol{\xi}}^\alpha$ is the difference between $\hat{\boldsymbol{\xi}}^\alpha$ and its expectation value

$$\boldsymbol{\xi} = \langle \hat{\boldsymbol{\xi}}^\alpha \rangle = \langle \hat{\boldsymbol{\xi}} \rangle \,. \tag{34}$$

Accordingly, if the assumption of independent sub-regions were true, the covariance of the total measurement $\hat{\boldsymbol{\xi}}$ could be estimated by

$$\hat{\mathbf{C}}_{\rm SC} = \frac{1}{N(N-1)}\sum_{\alpha=1}^{N} (\boldsymbol{\xi}^\alpha - \bar{\boldsymbol{\xi}})^{\rm T} \, (\boldsymbol{\xi}^\alpha - \bar{\boldsymbol{\xi}}) \,. \tag{35}$$

We will call the estimator in equation (35) the *sub-sample covariance* (Norberg et al. 2009). The main systematic effects of internal covariance estimation can be most easily understood in terms of this estimator. Hence, before introducing the jackknife and bootstrap estimator, we will explain these systematics in the following two sections.

4.2 Correlation of sub-samples

The sub-sample covariance estimator relies on the assumption that the data are split into independent sub-samples, i.e. that there is no correlation of the measurements of the correlation functions in the different sub-regions,

$$\langle \Delta\hat{\boldsymbol{\xi}}_i^\alpha \Delta\hat{\boldsymbol{\xi}}_j^\beta \rangle = 0 \,, \quad \text{for } \alpha \neq \beta \,. \tag{36}$$

This can be seen from the fact that equation (35) simply re-scales the sub-field-to-sub-field covariance by a factor of $1/N$ to get the covariance of the whole survey. If the sub-samples are correlated, this will underestimate the true covariance matrix (Nordman & Lahiri 2007).

Another way to think about this is as follows: the sub-sample covariance estimator assumes that the covariance matrix of $\hat{\boldsymbol{\xi}}$ is inversely proportional to the survey area A. Hence it estimates the covariance of sub-regions of the size $A_{\rm sub}$ within the data and then re-scales it to the total area,

$$\mathbf{C} = \frac{A_{\rm sub}}{A} \cdot \mathbf{C}_{\rm sub} = \frac{1}{N} \cdot \mathbf{C}_{\rm sub}, \tag{37}$$

where N is again the number of sub-regions. But already from the lognormal model for the covariance it can be seen, that this re-scaling is not valid. The lognormal correction term to the Gaussian covariance matrix is given by

$$\mathbf{C}_{\pm\pm}^{\rm ss,log}(\theta_1,\theta_2) = \frac{8\pi}{\kappa_0^2 A}\, \xi_\pm(\theta_1)\xi_\pm(\theta_2) \int_0^{\theta_A} {\rm d}\theta \; \theta \; \xi_+(\theta). \tag{38}$$

This term may be proportional to $1/A$, but the upper integral boundary also depends on the survey diameter θ_A. As A increases, the covariance therefore decreases slower than $1/A$. Hence, assuming $1/A$ scaling when extrapolating from the covariance of the smaller sub-fields to the covariance of the full area underestimates the full covariance. Also, note that even the Gaussian covariance term in equation (18) is only an approximation for large survey sizes A. It also suffers from a finite area effect as can be seen from its derivation in Schneider et al. (2002) or its form given in Hilbert et al. (2011).

The fact that sub-samples should be as uncorrelated as possible is also the reason why the re-sampling of the data should be done into spatially connected patches. If instead the data would be randomly divided into sub-samples then the shear correlations in the sub-samples would be almost identical. Hence, only the shape-noise contributions to the covariance would be measured by such an estimator.

4.3 Galaxy pairs crossing between sub-samples

A problem specific to the internal covariance estimation for two-point correlation functions is the question of what to do with pairs of galaxies where each galaxy lies in a different sub-region of the survey.

In fact, the pieces of information in a cosmic shear survey are not the individual galaxy shapes but the pairs of galaxy shapes. If the pairs crossing between sub-regions are completely ignored when computing the sub-measurements $\hat{\xi}^\alpha$, then one is re-sampling a data set that has *less* information than the total measurement of $\xi_\pm$ and hence a larger variance. Note, that this does not only influence the shape-noise part of the covariance but also the cosmic variance part. The reason is that galaxies at the edge of a sub-region contribute less terms to the correlation function measurement than galaxies in the centre of the sub-region (cf. Fig. 3), i.e. the area of the sub-patch is not uniformly probed by the galaxy pairs and the measured shear correlations are dominated by the inner part of the patch. In contrast to the correlation of sub-samples discussed before, this increases the cosmic variance between the sub-samples and can bias the covariance estimate high – especially on large angular scales.

This effect can in principle be resolved by re-sampling the set of pairs (instead of the set of galaxy shapes), i.e. by defining the sub-measurement $\hat{\boldsymbol{\xi}}^\alpha$ as

$$\hat{\xi}_\pm^\alpha(\theta) = \frac{\sum_{\text{pairs in }\alpha}(\epsilon_t^i\epsilon_t^j \pm \epsilon_\times^i\epsilon_\times^j) + \sum_{\text{half of cross pairs}}(\epsilon_t^i\epsilon_t^j \pm \epsilon_\times^i\epsilon_\times^j)}{N_{\rm pairs}}. \tag{39}$$

How this re-sampling of galaxy pairs can be done is illustrated in Fig. 4. Especially one has to make sure that each galaxy pair enters

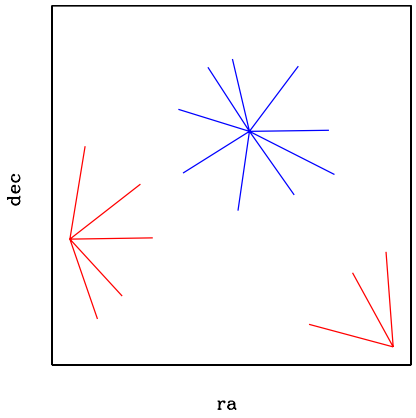

Figure 3. Galaxies at the edge of a sub-region (in red) contribute less pairs to the measurement of the correlation functions (i.e. to equation 5 applied to the sub-sample) than galaxies in the centre of the sub-region (in blue). Consequently, the area of the sub-patch is not uniformly probed by the galaxy pairs. This increases the cosmic variance between sub-regions and biases the covariance estimates high. Hence, it has an opposite effect than the correlation of sub-samples, which biases the covariance estimates low. As seen from the left-hand panel of Fig. 5, at large angular scales this can even lead to an overestimation of the cosmic variance of $\hat{\xi}_-$ (in the galaxy scheme, cf. also Fig. 4).

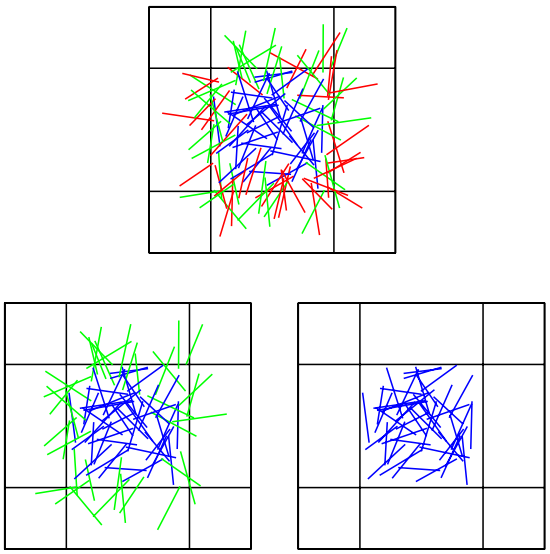

Figure 4. Two basic schemes for dividing a set of galaxy pairs into sub-samples: For each sub-region of the survey, there will be galaxy pairs crossing from that region into another (upper panel, green and red). In the galaxy scheme $\hat{\boldsymbol{\xi}}^\alpha$ is computed by considering only pairs that completely lay within the sub-region α (lower-right panel). In the pair–jackknife scheme (lower-left panel) half of the pairs that cross from α to another region (drawn in green) are taken into account for computing $\hat{\boldsymbol{\xi}}^\alpha$ while only the other half (red) is discarded.

exactly one of the $\hat{\boldsymbol{\xi}}^\alpha$. We call this procedure the *pair scheme*, while we will call the standard procedure of considering only the individual galaxies in sub-region α when computing $\hat{\boldsymbol{\xi}}^\alpha$ as *galaxy scheme*.

In Fig. 5, we demonstrate this effect along with the effect of correlated sub-samples that was discussed before. The left-hand panel shows sub-sample estimates of the variance of $\hat{\xi}_\pm$ in a simulated survey (corresponding to setup I in Table 1), where the shape noise was put to zero and where 400 sub-samples were used. Both the variance of $\hat{\xi}_+$ and $\hat{\xi}_-$ are severely underestimated on small scales, which is due to the correlation of sub-samples. At large angular scales, the galaxy scheme yields systematically higher value for the variance than the pair scheme and at least for $\hat{\xi}_-$ it can even overestimate the variance. This is due to the missing cross-pairs in the re-sampling.

For the right-hand panel of Fig. 5, we have generated a catalogue of pure shape noise ($\sigma_\epsilon, A, n_{\rm gal}$ as in setup I). This is the only situation where the assumption of uncorrelated sub-samples is valid. You can see that in this case the pair scheme is able to estimate the variance without bias. The galaxy scheme overestimates the variance for the reasons explained before. A downside of the pair scheme is that the shear signals in the sub-measurements $\hat{\boldsymbol{\xi}}^\alpha$ become even more correlated, as can also be seen from the left-hand panel of Fig. 5.

4.4 Jackknife

Another method of covariance estimation that Norberg et al. (2009) investigate is the *delete-one-jackknife*. Instead of estimating the covariance of the measurements $\boldsymbol{\xi}^\alpha$ and re-scaling it to the size of the whole survey the jackknife is considering the measurements

$$\hat{\xi}_\pm^{*\alpha}(\theta) = \frac{\sum_{\{i,j \text{ not in } \alpha\}} (\epsilon_t^i \epsilon_t^j \pm \epsilon_\times^i \epsilon_\times^j) \cdot \Delta_\theta(|\boldsymbol{\theta}_i - \boldsymbol{\theta}_j|)}{\sum_{\{i,j \text{ not in } \alpha\}} \Delta_\theta(|\boldsymbol{\theta}_i - \boldsymbol{\theta}_j|)}, \tag{40}$$

i.e. the *jackknife-sample* α is generated by cutting out the sub-region α and measuring the correlation functions in the rest of the survey. The jackknife estimate for the covariance matrix is then given by (Efron 1982; Norberg et al. 2009)

$$\hat{\mathbf{C}}_{\rm jack} = \frac{N-1}{N} \sum_{\alpha=1}^{N} (\boldsymbol{\xi}^{*\alpha} - \bar{\boldsymbol{\xi}}^*)^{\rm T} (\boldsymbol{\xi}^{*\alpha} - \bar{\boldsymbol{\xi}}^*), \tag{41}$$

where $\bar{\boldsymbol{\xi}}^*$ is the mean of all jackknife measurements.

If we again assume that all sub-regions have the same galaxy density and that the correlation functions are measured on scales much smaller than the sub-region size then $\boldsymbol{\xi}^{*\alpha}$ is approximately given by

$$\boldsymbol{\xi}^{*\alpha} \approx \frac{1}{N-1} \sum_{\beta \neq \alpha} \hat{\boldsymbol{\xi}}^\beta. \tag{42}$$

From this it also follows that

$$\begin{aligned}
\boldsymbol{\xi}^{*\alpha} - \bar{\boldsymbol{\xi}}^* &\approx \frac{1}{N-1} \sum_{\beta \neq \alpha} \hat{\boldsymbol{\xi}}^\beta - \frac{1}{N} \sum_\beta \boldsymbol{\xi}^{*\beta} \\
&= \frac{N \cdot \bar{\boldsymbol{\xi}} - \hat{\boldsymbol{\xi}}^\alpha}{N-1} - \frac{1}{(N-1) \cdot N} \sum_\beta \sum_{\gamma \neq \beta} \hat{\boldsymbol{\xi}}^\gamma \\
&= \frac{N \cdot \bar{\boldsymbol{\xi}} - \hat{\boldsymbol{\xi}}^\alpha}{N-1} - \frac{N-1}{(N-1) \cdot N} \sum_\gamma \hat{\boldsymbol{\xi}}^\gamma \\
&= \frac{N \cdot \bar{\boldsymbol{\xi}} - \hat{\boldsymbol{\xi}}^\alpha}{N-1} - \bar{\boldsymbol{\xi}} \\
&= \frac{\bar{\boldsymbol{\xi}} - \hat{\boldsymbol{\xi}}^\alpha}{N-1}.
\end{aligned} \tag{43}$$

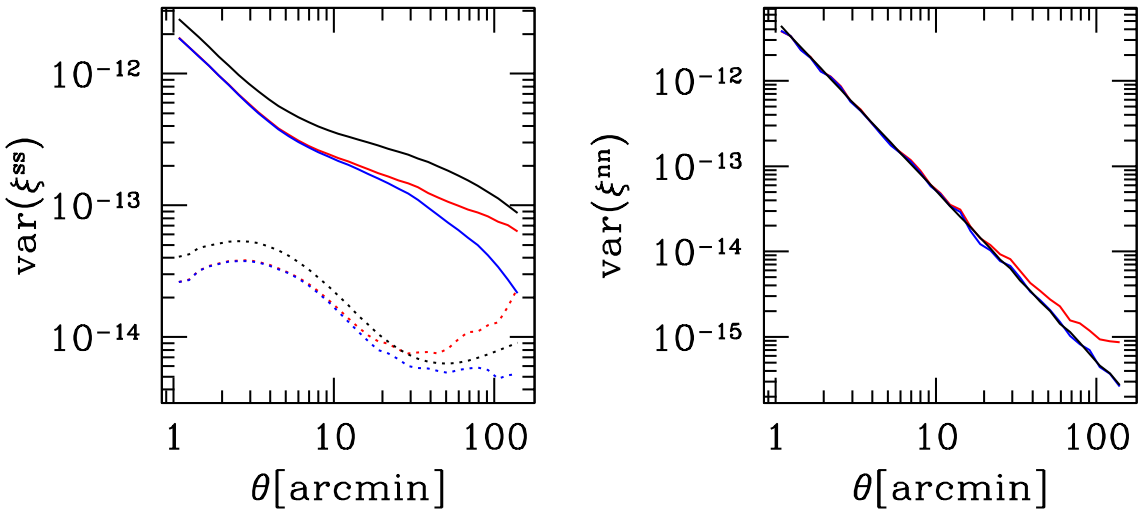

Figure 5. Different variance estimates using the sub-sample covariance estimator and 400 sub-samples. Left: variance estimates for $\hat{\boldsymbol{\xi}}_+$ (solid lines) and $\hat{\boldsymbol{\xi}}_-$ (dotted lines) in a mock catalogue without shape noise that is otherwise following setup I. The red lines show the galaxy scheme estimate (cf. Section 4.3), the blue lines show the pair scheme estimate and the black lines show the lognormal input model. Right: sub-sample estimates of the variance of $\hat{\boldsymbol{\xi}}_+$ in a mock catalogue that only consists of shape noise and has the same area and density as in setup I. It is only in this situation (and in the pair scheme) that internal estimation of the covariance yields unbiased results.

Inserting this into the definition of $\hat{\mathbf{C}}_{\rm jack}$ gives exactly the sub-sample covariance $\hat{\mathbf{C}}_{\rm SC}$, i.e. on small angular scales the two methods are approximately equivalent.[7]

In jackknife estimation one can in principle also differentiate between a pair scheme and a galaxy scheme. Using equation (40) for $\boldsymbol{\xi}^{*\alpha}$ corresponds to the galaxy scheme. This is equivalent to disregarding all pairs in the top panel of Fig. 4 when computing $\boldsymbol{\xi}^{*\alpha}$. The pair scheme is given by disregarding all pairs in the lower-left panel of Fig. 4 when computing $\boldsymbol{\xi}^{*\alpha}$. In the pair scheme jackknife and sub-sample covariance become exactly equivalent when assuming that each sub-patch has the same number of galaxies.

4.5 Bootstrap covariance

The so called *block bootstrap* estimator of the covariance also divides the data into sub-samples. If the data are split into N sub-regions, then a number of $N_{\rm boot}$ bootstrap re-samplings of the data are generated by randomly drawing with replacement N of the sub-samples and combining then into one re-sampled data set (Efron 1982; Nordman & Lahiri 2007; Loh 2008; Norberg et al. 2009). If the correlation function measured in the re-sampled data i ($i = 1, \ldots, N_{\rm boot}$) is called $\boldsymbol{\xi}^{{\rm boot},i}$, then the bootstrap estimate of the covariance is given by

$$\hat{\mathbf{C}}_{\rm boot} = \frac{1}{N_{\rm boot}-1}\sum_{i=1}^{N_{\rm boot}}(\boldsymbol{\xi}^{{\rm boot},i} - \bar{\boldsymbol{\xi}}^{\rm boot})^{\rm T}\,(\boldsymbol{\xi}^{{\rm boot},i} - \bar{\boldsymbol{\xi}}^{\rm boot}), \qquad (44)$$

where $\bar{\boldsymbol{\xi}}^{\rm boot}$ is now the mean of all $\boldsymbol{\xi}^{{\rm boot},i}$.

Again, the question arises of whether one should consider the single galaxies or the galaxy pairs as the actual data (cf. Section 4.3). In what we will call *galaxy–bootstrap* one simply adds a copy of all galaxies in a sub-region α to the re-sampled data set i each time the sub-region α gets drawn.

[7] This is no general statement on the jackknife method. It holds only in our particular situation.

In the *pair–bootstrap* one adds all pairs associated with sub-region α to the list of pairs that is used to compute $\boldsymbol{\xi}^{{\rm boot},i}$. The difference between the two bootstrap schemes is mainly the following: if the sub-region α gets drawn n times, then each pair in α gets a weight of n in the pair scheme and a weight of n^2 in the galaxy scheme.

Note that the pair–bootstrap is very similar to what Loh (2008) describes as *marked point bootstrap*, the only difference being, that we chose to split pairs between sub-regions evenly among these regions.

We will see in Section 5 that the galaxy–bootstrap severely overestimates the covariance. The other covariance estimators perform very similar to each other and suffer in similar ways from the systematics explained in Sections 4.2 and 4.3.

4.6 Stability and inversion of the covariance estimate

All effects that bias the internal covariance estimate can in principle be minimized by dividing the data into very large sub-regions. This decreases both the correlation of the different sub-regions and the influence of pairs crossing between sub-regions. However, this also decreases the possible number of re-samplings and hence increases the variance of the covariance estimator itself.

In order to derive constraints on the number of re-samplings let us assume that we are in the limit were the correlations between sub-regions are small. Small here means that

$$\langle\Delta\hat{\boldsymbol{\xi}}_i^{\alpha}\Delta\hat{\boldsymbol{\xi}}_j^{\beta}\rangle \ll \langle\Delta\hat{\boldsymbol{\xi}}_i^{\alpha}\Delta\hat{\boldsymbol{\xi}}_j^{\alpha}\rangle, \text{ for } \alpha \neq \beta. \qquad (45)$$

As explained before, this is the only limit in which internal covariance estimation is valid. In this limit, the sub-sample covariance is just a re-scaling of the sample covariance of independent realizations of the sub-regions. Hence – in the limit considered here and under the assumption that the data vector behaves Gaussian – the sub-sample covariance estimates are distributed according to a Wishart distribution (cf. Taylor et al. 2013). Also, the pair–jackknife is almost equivalent to the pair-version of the sub-sample covariance, i.e. to equation (35) when $\hat{\boldsymbol{\xi}}^{\alpha}$ is computed with equation (39).

Hence, also the pair–jackknife estimates should approximately follow a Wishart distribution.

The most important consequence of this is that the inverse of the covariance matrix estimate will be a biased estimate of the true inverse covariance matrix, and the bias is approximately given by (Hartlap, Simon & Schneider 2007; Taylor et al. 2013)

$$\langle \hat{\mathbf{C}}_{\rm SC}^{-1} \rangle \approx \frac{N-1}{N-d-2} \mathbf{C}_{\rm true}^{-1}, \tag{46}$$

where N is the number of sub-regions and d is the number of data points in $\hat{\boldsymbol{\xi}}$. Especially, this factor has to be accounted for when computing the χ^2 statistic, equation (31), i.e. it has an influence on the constraints derived on cosmological parameters when using internal covariance estimation.

Taylor et al. (2013) also give constraints on N with respect to d when a certain accuracy in the final parameter constraints is demanded.[8] There derivation, however, assumes an exact Wishart distribution. Hence, it only applies to internal covariance estimates if the data vector is Gaussian and if the sub-regions are large enough to be independent. Nevertheless, we take their criterion,

$$N \stackrel{!}{>} \frac{2}{\epsilon^2} + (d+4), \tag{47}$$

where ϵ is the required fractional accuracy on parameter constraints, as a guideline also for internal covariance estimators. Demanding a fractional accuracy of $\epsilon = 0.2$ for the parameter constraints, this yields a necessary number of $N > 54 + d$ re-sampling. Below this number there is no chance for internal covariance estimators to yield parameter uncertainties that are accurate to more than 20 per cent.

5 TESTING INTERNAL COVARIANCE ESTIMATORS ON SIMULATED COSMIC SHEAR SURVEYS

We will now use the simulations described in Section 3 to test the performance of internal covariance estimators. First, we will use setup I (cf. 1) corresponding to a rather deep survey. We carry out 50 independent realizations of this survey. In each survey, we measure the correlation functions in the range and binning that was explained in Section 3. We then estimate the covariance of the measured correlation functions using the different internal estimation schemes that were introduced in Section 4. Throughout this section – except for Section 5.2 – we consider the lognormal model that was explained in Section 2.2 as the *true* covariance of the simulated surveys. This is justified by the fact that our results do not change if we instead use the sample covariance of 1000 independent realizations that were presented in Section 3.

In Fig. 6, we compare the sub-sample, jackknife and bootstrap estimates of the diagonal elements of the covariance matrix (both in the galaxy and pair scheme) when splitting the survey into $N = 225$ sub-regions. The most impressive finding is, that in the galaxy scheme the bootstrap severely overestimates the variance. This is in agreement with the findings of Norberg et al. (2009) for galaxy clustering correlation functions. The duplication of whole sub-volumes of galaxies creates bootstrap samples that are in fact unrealistic, i.e. these bootstrap samples contain regions with no sources at all and on the other hand regions with a very high source density. Each

[8] However, they are ignoring the impact that the variance in the inverted covariance estimate has on parameter constraints, which is investigated by Taylor & Joachimi (2014).

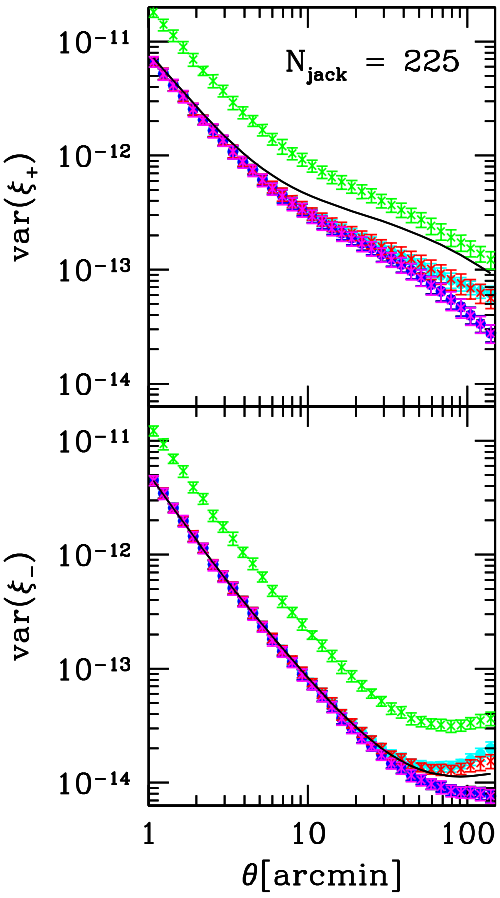

Figure 6. A comparison of the different internal estimation schemes when splitting the survey into $N = 225$ sub-regions. Green: galaxy–bootstrap, purple: pair–bootstrap, red: galaxy–jackknife, blue: pair–jackknife and cyan: sub-sample covariance compared to the analytical covariance (black line). We show the sub-sample covariance only in the galaxy scheme because in the pair scheme it is almost identical to jackknife and bootstrap. As explained in Section 4, at large angular scales the different treatment of galaxy pairs crossing between sub-region leads to an overestimation of the variance by the galaxy scheme and an underestimation of the variance by the pair scheme.

original galaxy pair gets weighted by a factor of n^2 when the corresponding region is drawn n times. This puts a very high weight on small sub-areas of the bootstrap sample and creates an unphysically high variance between the bootstrap samples.

In the pair scheme however, all three internal estimators perform almost identical. This is not surprising, because in that scheme the bootstrap is just an approximation to the sub-sample covariance and sub-sample and jackknife covariance are almost identical in the pair scheme. As explained in Section 4.3, in the galaxy–jackknife scheme the two effects of correlated sub-regions and false re-sampling of pairs partly cancel each other. Hence the galaxy–jackknife comes closest to the true variance at large scales. The performance of the sub-sample covariance (in the galaxy scheme) only slightly differs from that.

Because of the strong similarity between the different estimators, we will restrict the following analyses to the pair–jackknife and the galaxy–jackknife. We now investigate the influence of sub-region size on internal covariance estimation. Hence we split the surveys into three different numbers of sub-regions: 10^2, 15^2 and 20^2 corresponding to sub-region areas of approximately 7.0×7.0,

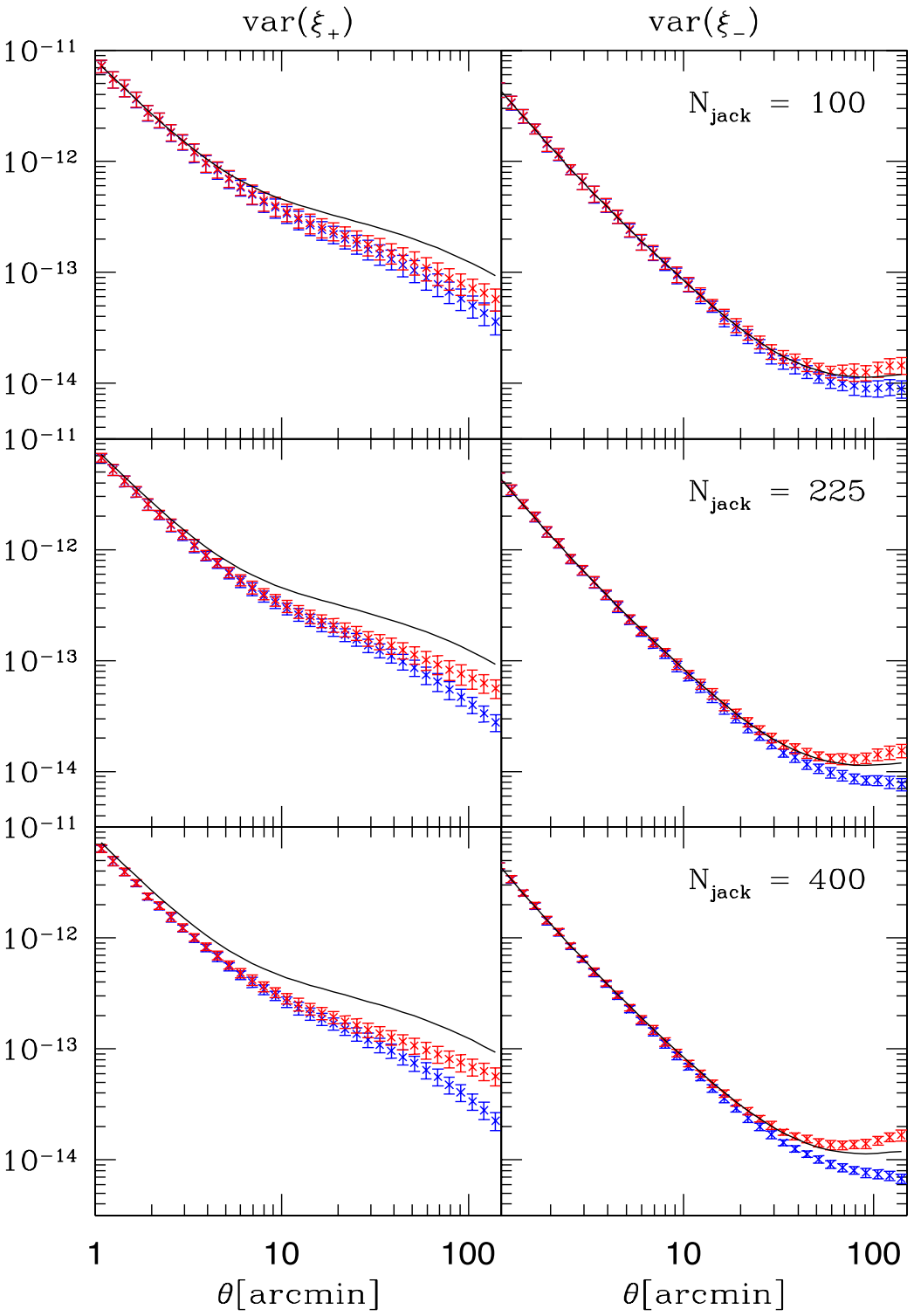

Figure 7. Mean values of 50 jackknife estimates of the variance of ξ_+ (left) and ξ_- (right). Galaxy–jackknife was used for the red points, while pair–jackknife was used for the blue points and the error bars show the standard deviation in the 50 mock surveys). The black line corresponds to the lognormal input model of the simulations.

4.67×4.67 and 3.5×3.5 deg^2. In Fig. 7, we compare the mean value of the 50 jackknife estimates of the variance of $\hat{\xi}_\pm$ (the diagonal elements of the covariance matrix) to the true underlying lognormal model. A comparison of the off-diagonal behaviour of the jackknife estimates to that of the input-covariance can be found in Appendix A. The error bars in Fig. 7 represent the standard deviation of the 50 jackknife estimates, i.e. they illustrate the noise of the internal estimators. You can see in this figure the biases in the jackknife estimates that were explained in the previous section. For ξ_+, both jackknife schemes underestimate the variance. At large scales, this is in the galaxy–jackknife scheme partly compensated by the false re-sampling of galaxy pairs. For ξ_-, the pair–jackknife

underestimates the variance while the galaxy–jackknife overestimates it. ξ_- is a much more local measure in the sense that the different sub-regions are less correlated in $\hat{\xi}_-$ and that the covariance matrix is much more dominated by the shape noise contributions. Hence, the severe systematic underestimation of the variance that can be seen for ξ_+ does not appear as strongly for ξ_-.

When increasing the number of sub-regions for the jackknife estimators, the noise in the variance estimates becomes smaller but the deviations from the true variance also become stronger. This is because for smaller sub-regions the estimated $\hat{\boldsymbol{\xi}}^{\alpha}$ become more correlated and because there will be more galaxy pairs crossing from one sub-region to another.

5.1 Constraints on cosmological parameters

We will now take the 50 simulations as mock observations and try to constrain the dark matter density parameter Ω_m and the power spectrum normalization σ_8. To do so, we sample the Ω_m–σ_8 plane on a fine grid while keeping the other cosmological parameters fixed. Following a Bayesian approach, we take the probability density in the parameter space to be proportional to the likelihood,

$$p(\boldsymbol{\pi}) \sim \mathcal{L}(\boldsymbol{\pi}) \sim \exp\left(-\frac{1}{2}\chi^2[\boldsymbol{\pi}]\right), \tag{48}$$

where we assume our data vector $\hat{\boldsymbol{\xi}}$ to be Gaussian such that

$$\chi^2[\boldsymbol{\pi}] = (\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}[\boldsymbol{\pi}])^\mathrm{T}\mathbf{C}^{-1}(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}[\boldsymbol{\pi}]). \tag{49}$$

Here, $\boldsymbol{\xi}[\boldsymbol{\pi}]$ are our model predictions for $\langle\hat{\boldsymbol{\xi}}\rangle$ which we again compute with the NICAEA package. We are assuming a prior of $\Omega_\mathrm{m} \in [0.1, 0.4]$ and $\sigma_8 \in [0.8, 1.1]$, which is well centred around our input cosmology. For $\mathbf{C}$ we will either insert the lognormal model covariance or the jackknife estimates of the covariance. We will de-bias the inverse of the latter in the way explained in Section 4.6. Note that the reasoning in Section 4.6 is in principle only valid for the pair–jackknife. And also for the pair–jackknife it is only valid in the case of almost uncorrelated sub-regions. We will nevertheless carry out the de-biasing in the same way for both jackknife schemes. Furthermore, we will also ignore the variance of the inverted covariance estimate (Taylor & Joachimi 2014), as explained in the end of Section 4. Our data vector $\hat{\boldsymbol{\xi}}$ will be either $\hat{\boldsymbol{\xi}}_+$ or $\hat{\boldsymbol{\xi}}_-$ or the joint data vector of both correlation functions, in which case we will also take into account the cross-covariance between the two.

For each mock observation $\hat{\boldsymbol{\xi}}$ and for each available covariance matrix, we use equation (48) to compute marginalized 1σ constraints on Ω_m and σ_8, i.e. we consider the marginalized probability densities

$$p_\Omega(\Omega_\mathrm{m}) = \int \mathrm{d}\sigma_8\ p(\Omega_\mathrm{m}, \sigma_8)$$

$$p_\sigma(\sigma_8) = \int \mathrm{d}\Omega_\mathrm{m}\ p(\Omega_\mathrm{m}, \sigma_8) \tag{50}$$

and we define 1σ confidence interval to be that interval around the best-fitting parameter value which encloses ∼68 per cent of the probability and which has equal values of the probability density at each interval boundary.[9]

Because of the strong degeneracy between Ω_m and σ_8 (Kilbinger & Schneider 2004; Kilbinger et al. 2013), even little uncertainties in the modelling of $\boldsymbol{\xi}[\boldsymbol{\pi}]$[10] or in our simulations could shift the best-fitting values of the parameters along the degeneracy. Fortunately, this does not affect our analysis because we only have to compare the constraints derived from the jackknife covariance estimates to the constraints obtained from the true (lognormal) covariance matrix. Furthermore, our results do not change noticeably, if instead of the lognormal covariance matrix we use the sample covariance estimated from 1000 simulations (cf. Section 3). Hence in any case, our analysis provides a fair test of internal covariance estimators.

In Fig. 8, we show the mean values of the upper and lower boundaries on Ω_m and σ_8 as well as their mean best-fitting value for different numbers of jackknife re-samplings (red points and error bars). The mean is taken with respect to all 50 confidence intervals we computed from the 50 mock observations. We also compare the jackknife constraints to those we get when using the true covariance matrix (blue lines). These figures only show the results for the galaxy–jackknife, which in the situation considered here yields the best agreement with the true covariance.

We compare galaxy–jackknife and pair–jackknife in Fig. 9. Here we show the mean width of the confidence intervals obtained with galaxy–jackknife, pair–jackknife and the true covariance matrix. For ξ_-, the width of the confidence intervals agrees well with the confidence intervals obtained from the true covariance matrix. This is because the covariance matrix of ξ_- is dominated by its shape noise component, which is very accurately captured by jackknife. In fact, even for the pair scheme and even for 400 jackknife re-samplings the width of the confidence intervals from ξ_- alone is not underestimated. This seems to contradict Fig. 7, where the pair scheme systematically underestimates the covariance. One reason for this is probably, that the variance in the inverted covariance estimate increases parameter uncertainties (Taylor & Joachimi 2014). Note especially, that this is not the same effect as the de-biasing in equation (4.6). For ξ_+, the strong underestimation of the covariance matrix by jackknife also leads to an underestimation of the uncertainties on Ω_m and σ_8. Again one can see that the variance in the width of the confidence intervals (the error bars in Fig. 9) becomes smaller, when more jackknife re-samplings are used. In turn, this increases the overall underestimation of the uncertainties. If both correlation functions are combined and 225 re-samplings are used, the parameter uncertainties are underestimated by ∼10 per cent.

We have not shown results from the pair–jackknife estimates in Fig. 8, but the best-fitting values of Ω_m and σ_8 agree very well between the two jackknife schemes (i.e. within the green error bars in Fig. 8), if only $\hat{\boldsymbol{\xi}}_+$ or $\hat{\boldsymbol{\xi}}_-$ are used to constrain the parameters. In Fig. 10, we compare the pair–jackknife and galaxy–jackknife best-fitting values when using the *full* data vector. Here the pair–jackknife seems to yield a stronger bias of the best-fitting values with respect to the true covariance.

The above results indicate that internal covariance estimation can reproduce the constraints on parameters from the true covariance quite well, especially when the galaxy–jackknife scheme is used. However, these results are not generalizable. In general, internal estimation of the covariance works best if the covariance matrix is shape noise dominated. Hence, the answer to what is the best estimation scheme and how well it can reproduce the true error bars on cosmological parameters depends on the depth of the considered survey. A shallower survey not only has a smaller source density and hence a bigger shape noise. It also has a smaller convergence

[9] Without the last statement the definition of the 1σ confidence interval would be ambiguous.

[10] In our modelling, we are for example not considering the finite bin width in our measurement of $\hat{\boldsymbol{\xi}}$.

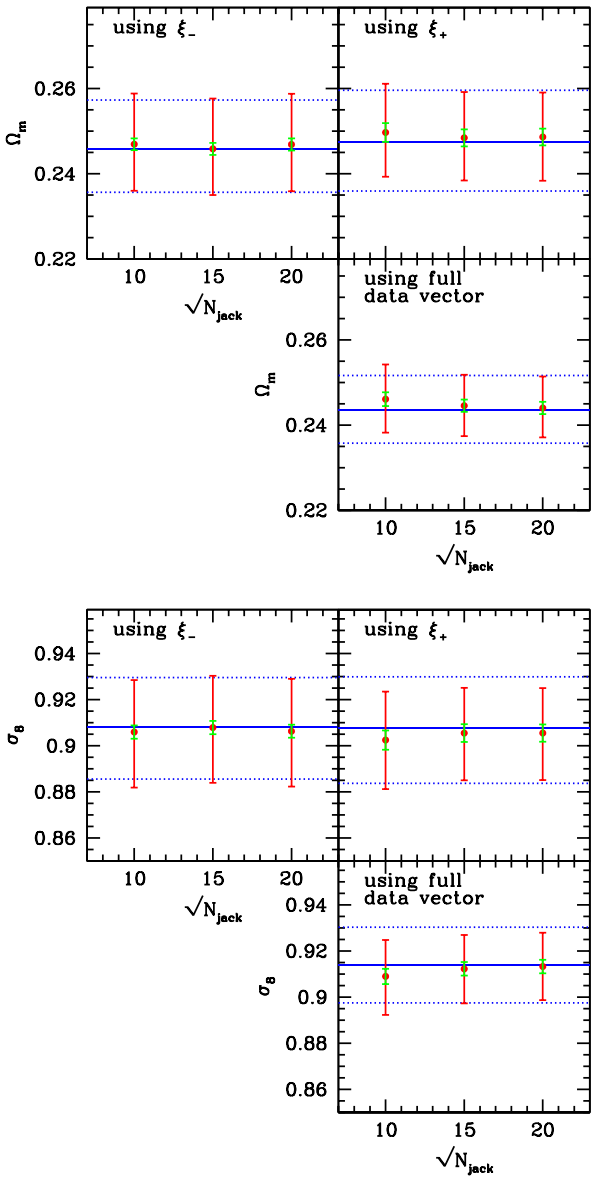

Figure 8. Mean 1σ constraints on Ω_m and σ_8 using galaxy–jackknife (red error bars). The green error bars show the standard deviation of the mean best-fitting values (i.e. the standard deviation of the best-fitting values divided by $\sqrt{50}$). The blue lines indicate the constraints that are obtained when the true covariance is used in each mock catalogue. Note that the error bars are very symmetric. For surveys as big as our simulations the constraining power becomes large enough to turn the – usually banana shaped – degeneracy between Ω_m and σ_8 into almost elliptical contours in the parameter plane (cf. Appendix B).

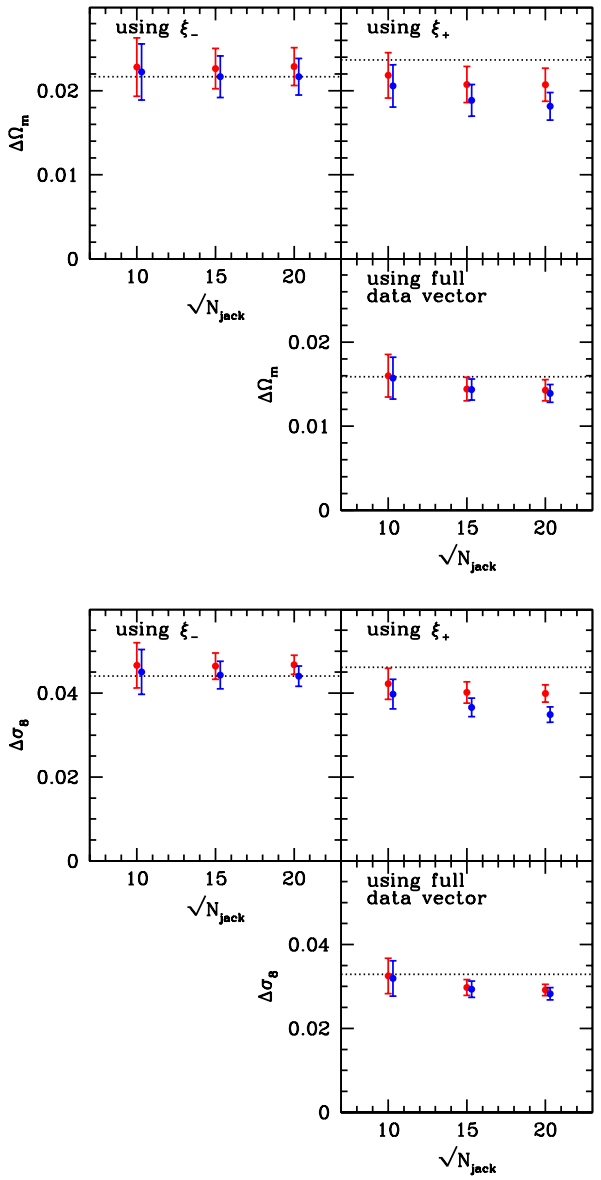

Figure 9. Mean width of the 1σ uncertainty on Ω_m and σ_8 using pair–jackknife (blue) and galaxy–jackknife (red). The error bars show the standard deviation of the 50 estimated confidence intervals. The black dotted line indicates the mean width of the confidence intervals when the true covariance is used in each mock catalogue.

power spectrum which in turn reduces the cosmic variance part of the covariance.

The procedure we presented above to investigate the performance of internal covariance estimators thus has to be re-run for each survey under consideration. One can consider the lognormal model as a good model for the true covariance of our simulations for mock catalogues with an area of $\gtrsim$1000 deg^2 and a simple, connected geometry. For smaller surveys the finite-area-effect should not be ignored (Sato et al. 2011; Kilbinger et al. 2013). However, these surveys can be simulated fast enough with our public code to generate a large sample of independent realizations of the mock data which provides a good sample covariance estimate of the true covariance matrix. This estimate can then be compared to an ensemble of internal covariance estimates as we have done it above.

5.2 Matching the procedure to DES-SV and year 5 data

We will now present an application of our method. Our attempt is to determine the performance of internal covariance estimation for

(i) setup IIa: Dark Energy Survey science verification data (DES-SV)
(ii) setup IIb: DES year five data (DES-Y5) assuming a low source density
(iii) setup IIc: DES year five data assuming a high source density.

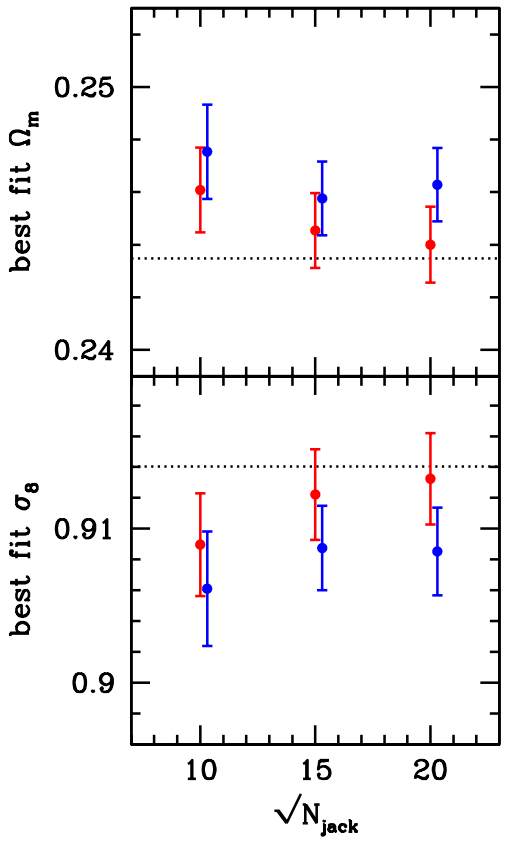

Figure 10. Mean best-fitting values of Ω_m and σ_8 using pair–jackknife (blue) and galaxy–jackknife (red). The error bars show the standard deviation of the mean, as estimated from the 50 best-fitting values. The black dotted line indicates the mean best-fitting value when the true covariance is used in each mock catalogue.

For the area, shape noise, source density and source redshift distribution cf. Table 1 and Section 3.1. A source density of 10 arcmin^{-2} is forecasted for the final DES data, while a density of $\sim$6 arcmin^{-2} roughly corresponds to the current status of DES-SV data. Note also, that we are using a mask similar to the footprint of DES-SV to simulate mock shape catalogues for setup IIa. Setups IIb and IIc are simply simulated to be square shaped.

We adjust our data vector to that used by Becker et al. (2015), i.e. for both ξ_+ and ξ_- we now use 15 logarithmic bins ranging from $\theta = 2$ arcmin to $\theta = 300$ arcmin. We will cut the survey into 100 sub-regions for setup IIa. Note that this way our biggest angular scales by far exceed the diameter of our sub-regions which is $\sim$45 arcmin. Hence, this can be considered an on-the-edge test of internal covariance estimators. A good tool to define sub-regions in an irregular survey geometry is the *kmeans* algorithm.[11] For setups IIb and IIc, we decide to split the survey into 225 sub-regions which corresponds to a diameter of $\sim 4.7^\circ$. This should give a more stable estimate of the covariance while still yielding much larger sub-regions than in setup IIa. In Fig. 11, we compare the internal variance estimates to the true covariance. The latter is taken to be the lognormal model for the Y5 simulations and a sample variance computed from 1000 independent realizations for the SV simulations. Because of the fewer number of bins we are now using the procedure described in Section 2.2.3 to compute the lognormal covariance matrix. As you can see, for $\hat{\xi}_-$ the pair–jackknife now becomes the best estimator of the variance. For $\hat{\xi}_+$ the situation is similar to what we have seen before, i.e. both schemes mostly underestimate the variance and the galaxy–jackknife is overall closer to the true variance. Hence, judging from Fig. 11 we conclude that galaxy–jackknife should be used in order to not underestimate the true uncertainties in the data vector. However, these statements only hold for the diagonal elements of the covariance matrix. A convenient way to compare the complete covariance estimates is to derive likelihood contours from them in the desired parameter space.

We carry out a likelihood analysis in the Ω_m–σ_8 plane for the 10 simulations that have a Y5-like area and a source density of 6 arcmin^{-2} which is the highest density currently achieved in DES-SV data (Becker et al. 2015). In Fig. 12, we show the likelihood contours obtained from one of the simulations when using galaxy–jackknife, pair–jackknife and the lognormal model for the covariance matrix. The contours were obtained from Monte Carlo-Markov-Chains (MCMC, 150.000 steps) using the COSMOLIKE package by Eifler et al. (2014). We present the likelihood contours from the other nine independent simulations in Appendix B. As expected, jackknife estimation underestimates the uncertainties. The input cosmology lies within the 1σ contour in six of 10 simulation, when the lognormal covariance is used. It lies within the 1σ contour in five of 10 simulation, when the covariance is estimated with jackknife (either scheme).

In Table 2, we show the average ratio of the volume in the Ω_m–σ_8 plane enclosed by the 1σ and 2σ contours when using jackknife to that when using the true covariance matrix. Since the 1σ and 2σ ellipses obtained from jackknife and from the true covariance lie well on top of each other, this ratio can be considered as the fraction of the true uncertainties that is recovered by the jackknife covariance matrices. You can see from Table 2 that the volume inside contours of constant likelihood in the Ω_m–σ_8 plane estimated with galaxy–jackknife is on average $\gtrsim$85 per cent of the true volume, while the volume estimated with pair–jackknife recovers only $\gtrsim$70 per cent of the true volume. This agrees with the impression (from Figs B1 and B2) that the contours obtained with galaxy–jackknife match better to the contours obtained from the true covariance. Note also, that the ellipses obtained from pair–jackknife have in some cases a strong off-set along the degeneracy between Ω_m and σ_8 compared to the true covariance and the galaxy–jackknife estimates. This is probably because pair–jackknife strongly underestimates the variance of $\hat{\xi}_\pm$ at large angular scales, which causes even small fluctuations at these scales to shift the contours considerably.

Finally, we want to see how well jackknife matrices recover the uncertainties perpendicular to the degeneracy between Ω_m and σ_8. To do so, we consider the parameter combination

$$\Sigma_8 := \frac{\sigma_8}{0.9}\left(\frac{\Omega_m}{0.25}\right)^{0.5}. \quad (51)$$

Contours of constant Σ_8 are roughly parallel to the degeneracy that can be seen in Figs 12, B1 and B2. For each of our 10 realizations, we bin our MCMC's in Σ_8 to estimate its probability density. Table 3 displays the average ratio of the 1σ and 2σ uncertainties obtained from jackknife to the uncertainties obtained from the true covariance. This time, we find that galaxy–jackknife on average yields $\sim$90 per cent of the true uncertainties, while pair–jackknife yields $\sim$85 per cent. Hence, when the degeneracy between Ω_m and σ_8 is broken by other probes (such as the power spectrum of temperature fluctuation in the CMB) the performance of jackknife covariance matrices slightly improves.

[11] Implemented by Erin Sheldon for PYTHON, www.github.com/esheldon/kmeans_radec

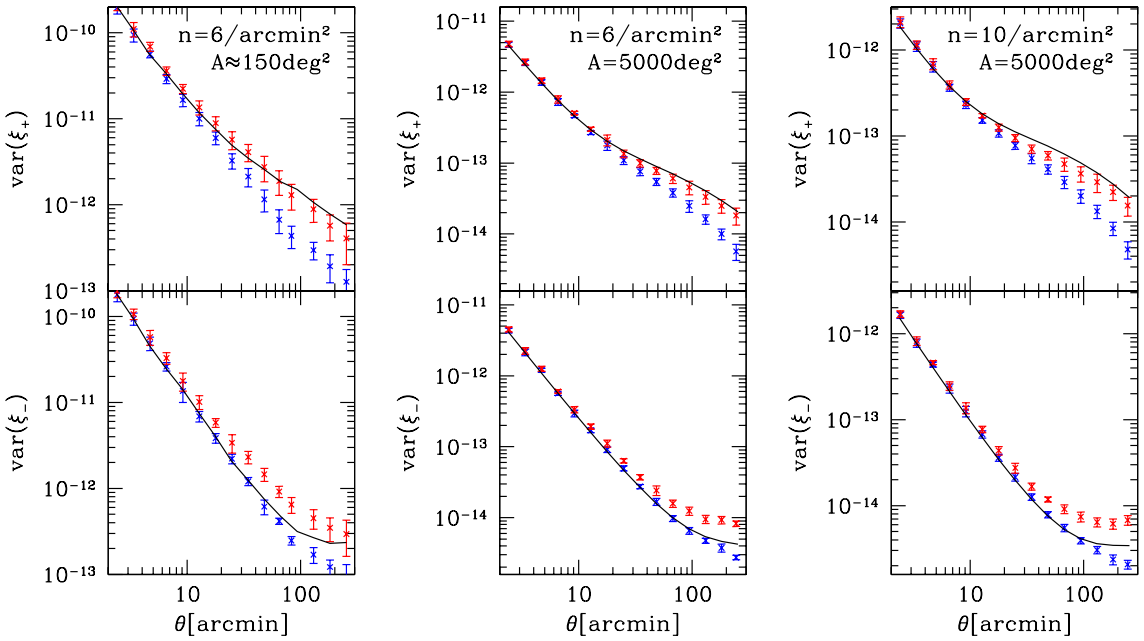

Figure 11. Variance estimates for DES-SV like data (left), DES-Y5 like data with a low density (middle) and with a high density (right). Red dots show the galaxy–jackknife estimates and blue dots the pair–jackknife estimates. For the Y5 case the lognormal model together with equation (20) was taken as a reference covariance (black lines), while for the SV case we estimated the true covariance from 1000 independent realizations of the mock data in order to account for the finite-area-effect. The error bars indicate the standard deviation of the single estimates as obtained from 10 independent measurements.

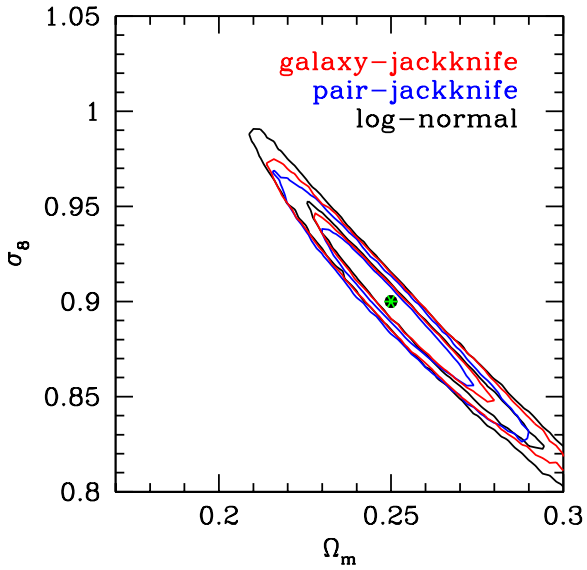

Figure 12. 1σ and 2σ contours in the Ω_m–σ_8 plane obtained from the two jackknife schemes (read and blue) and the true covariance (lognormal covariance, black) and using the combined data vector $(\hat{\xi}_+, \hat{\xi}_-)$. The input cosmology lies within the 1σ contour in six of 10 simulation, when the lognormal covariance is used. It lies within the 1σ contour in five of 10 simulation, when the covariance is estimated with jackknife (either scheme). In Appendix B, we show the contours obtained from the other simulations. The underestimation of the uncertainties by jackknife mainly takes place along the direction of the degeneracy between Ω_m and σ_8.

Judging from the above numbers and from the contours in Appendix B, we conclude that $\gtrsim$85 per cent of the true uncertainties on Ω_m and σ_8 in a 2D cosmic shear analysis can be recovered without the use of large suits of *N*-body simulations or covariance models. When other probes like the CMB are used to break the degeneracy between the two parameters, the performance of jackknife even increases, because the deviations from the true covariance mostly take place along the direction of degeneracy between Ω_m and σ_8.

Table 2. Ratio of the volume within the 1σ and 2σ contours in the Ω_m–σ_8 plane obtained from jackknife and true covariance (setup IIb). The errors are given by the standard deviation of a sample of 10 independent simulations. The combined data vector of ξ_+ and ξ_- was used.

	Galaxy–jackknife	Pair–jackknife
$V_{1\sigma,\,\mathrm{jack}}/V_{1\sigma,\,\mathrm{true}}$	0.86 ± 0.08	0.72 ± 0.09
$V_{2\sigma,\,\mathrm{jack}}/V_{2\sigma,\,\mathrm{true}}$	0.87 ± 0.08	0.74 ± 0.09

Table 3. Ratio of the 1σ and 2σ uncertainties on $\Sigma_8 \sim \sigma_8\Omega_m^{0.5}$ obtained from jackknife and true covariance (setup IIb). The errors are given by the standard deviation of a sample of 10 independent simulations. Again, the combined data vector of ξ_+ and ξ_- was used.

	Galaxy–jackknife	Pair–jackknife
$\Delta\Sigma_{8\,1\sigma,\,\mathrm{jack}}/\Delta\Sigma_{8\,1\sigma,\,\mathrm{true}}$	0.91 ± 0.08	0.86 ± 0.10
$\Delta\Sigma_{8\,2\sigma,\,\mathrm{jack}}/\Delta\Sigma_{8\,2\sigma,\,\mathrm{true}}$	0.90 ± 0.08	0.85 ± 0.09

6 CONCLUSIONS

We have explored the performance of internal covariance estimation for cosmic shear two-point correlation functions. We devised two different jackknife schemes and explained in detail when these schemes underestimate the true covariance and when overestimation takes place. Furthermore, we explained why the subsample covariance and the pair–bootstrap covariance yield results that are very similar to jackknife estimation of the covariance matrix. Based on the pair–jackknife scheme, we have argued that the Anderson-Hartlap-Kaufman (Kaufman 1967; Hartlap et al. 2007) de-biasing factor should also be applied when inverting jackknife

covariance matrices. Based on empirical findings, we also recommend the use of this factor for the galaxy–jackknife scheme.

We have demonstrated our findings in an exemplary study using lognormal simulations of the convergence field and the corresponding shear field. We found the performance of all internal covariance estimators – except for the bootstrapping of galaxies – to be very similar. For the investigated cases, jackknife covariance matrices can provide accurate uncertainties on cosmological parameters. Our conclusions regarding the two possible re-sampling schemes are the following.

(i) Galaxy–bootstrap severely overestimates the covariance, which is in agreement with the finding of Norberg et al. (2009) for galaxy clustering correlation functions.
(ii) From ξ_- alone, the pair–jackknife scheme reconstructs the parameter constraints most faithfully (cf. Fig. 8).
(iii) From ξ_+ alone and when combining the two correlation functions, we find that the parameter constraints are best reconstructed by the galaxy–jackknife.

The performance of the galaxy scheme turns out to be better in most situations, because its two systematic errors (cf. Sections 4.2 and 4.3) cancel each other partly. The pair–jackknife suffers from only one of these systematics and hence always yields lower (absolute) values for the covariance than the galaxy–jackknife and always underestimates the (absolute) values of the true covariance matrix.

Our results cannot be generalized to arbitrary surveys, i.e. our paper rather demonstrates a general method to find a good covariance estimation scheme for any particular survey. In making our simulation code public, we provide our readers with a tool to re-do the presented analyses for their desired setup. As an application example, we tested jackknife estimation of the covariance for a 2D cosmic shear analysis of the DES. We found that for the complete, 5-yr DES survey internal covariance estimators can provide reliable parameter constraints in a 2D cosmic shear analysis. We recommend a scheme of $\sim 15 \times 15$ jackknife re-samplings to yield a stable covariance matrix. Judging from Figs 12, B1 and B2, we find as before that the likelihood contours in the Ω_m–σ_8 plane are best reconstructed by the galaxy–jackknife scheme, if both correlation functions ξ_+ and ξ_- are combined. This way, on average $\gtrsim$85 per cent of the true uncertainties are captured by the internally estimated covariance matrix. If the degeneracy between Ω_m and σ_8 is broken, this value increases to $\sim$90 per cent. Hence, up to $\sim$90 per cent of the true uncertainties in a 2D cosmic shear analysis can be provided from internally estimated covariance matrices.

ACKNOWLEDGEMENTS

This work was supported by SFB-Transregio 33 'The Dark Universe' by the Deutsche Forschungsgemeinschaft (DFG). We also acknowledge the support by the DFG Cluster of Excellence 'Origin and Structure of the Universe'. The simulations have been carried out on the computing facilities of the Computational Center for Particle and Astrophysics (C2PAP). Part of the research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

This paper has gone through internal review by the DES collaboration. We thank David Bacon, Gary Berstein, Stefan Hilbert, Klaus Honscheid, Benjamin Joachimi and Bhuvnesh Jain for very helpful comments and discussions during the review process. Also, we would like to thank the anonymous referee for helpful suggestions on our manuscript.

Funding for the DES projects has been provided by the US Department of Energy, the US National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, the Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da Ciência, Tecnologia e Inovação, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the DES. The DES data management system is supported by the National Science Foundation under Grant Number AST-1138766.

The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Enérgeticas, Medioambientales y Tecnológicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l'Espai (IEEC/CSIC), the Institut de Física d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, the National Optical Astronomy Observatory, the University of Nottingham, The Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, and Texas A&M University.

The DES participants from Spanish institutions are partially supported by MINECO under grants AYA2012-39559, ESP2013-48274, FPA2013-47986 and Centro de Excelencia Severo Ochoa SEV-2012-0234. Research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) including ERC grant agreements 240672, 291329 and 306478.

REFERENCES

Becker M. R., Troxel M. A., MacCrann N., Krause E., Eifler T. F., Friedrich O., Nicola A., The DES Collaboration, 2015, preprint (arXiv:e-prints)
Cabré A., Fosalba P., Gaztañaga E., Manera M., 2007, MNRAS, 381, 1347
Cooray A., Hu W., 2001, ApJ, 554, 56
Crocce M., Cabré A., Gaztañaga E., 2011, MNRAS, 414, 329
de Simoni F. et al., 2013, MNRAS, 435, 3017
Efron B., 1982, The Jackknife, the Bootstrap and other Resampling Plans, CBMS-NSF Regional Conf. Ser. Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia
Eifler T., Schneider P., Hartlap J., 2009, A&A, 502, 721
Eifler T., Krause E., Schneider P., Honscheid K., 2014, MNRAS, 440, 1379
Flaugher B., 2005, Int. J. Mod. Phys. A, 20, 3121
Harnois-Déraps J., van Waerbeke L., 2015, MNRAS, 450, 2857
Hartlap J., Simon P., Schneider P., 2007, A&A, 464, 399
Hilbert S., Hartlap J., Schneider P., 2011, A&A, 536, A85
Jarvis M., Bernstein G., Jain B., 2004, MNRAS, 352, 338
Joachimi B., Schneider P., Eifler T., 2008, A&A, 477, 43
Kaiser N., Squires G., 1993, ApJ, 404, 441

Kaufman G. M., 1967, Report No. 6710, Center for Operations Research and Econometrics, Catholic Univ. Louvain, Heverlee, Belgium
Kilbinger M., Schneider P., 2004, A&A, 413, 465
Kilbinger M. et al., 2009, A&A, 497, 677
Kilbinger M. et al., 2013, MNRAS, 430, 2200
Loh J. M., 2008, ApJ, 681, 726
Laureijs R. et al., 2011, (arXiv:e-prints)
Martin S., Schneider P., Simon P., 2012, A&A, 540, A9
Norberg P., Baugh C. M., Gaztañaga E., Croton D. J., 2009, MNRAS, 396, 19
Nordman D. J., Lahiri S. N., 2007, Indian J. Stat., 69, 468
Sato M., Hamana T., Takahashi R., Takada M., Yoshida N., Matsubara T., Sugiyama N., 2009, ApJ, 701, 945
Sato M., Takada M., Hamana T., Matsubara T., 2011, ApJ, 734, 76
Schneider P., van Waerbeke L., Kilbinger M., Mellier Y., 2002, A&A, 396, 1
Semboloni E., van Waerbeke L., Heymans C., Hamana T., Colombi S., White M., Mellier Y., 2007, MNRAS, 375, L6
Simon P., King L. J., Schneider P., 2004, A&A, 417, 873
Smith R. E. et al., 2003, MNRAS, 341, 1311
Takada M., Jain B., 2009, MNRAS, 395, 2065
Takahashi R. et al., 2009, ApJ, 700, 479
Takahashi R., Soma S., Takada M., Kayo I., 2014, MNRAS, 444, 3473
Taruya A., Takada M., Hamana T., Kayo I., Futamase T., 2002, ApJ, 571, 638
Taylor A., Joachimi B., 2014, MNRAS, 442, 2728
Taylor A., Joachimi B., Kitching T., 2013, MNRAS, 432, 1928
The Dark Energy Survey Collaboration, 2005, (arXiv:e-prints)
Thomas S., Abdalla F., Lahav O., 2011, MNRAS, 412, 1669
Vale C., White M., 2003, ApJ, 592, 699
Wang Y., Brunner R. J., Dolence J. C., 2013, MNRAS, 432, 1961

APPENDIX A: CORRELATION MATRICES AND CONSTRAINTS FROM EMPIRICAL COVARIANCE

To see how jackknife estimates of the covariance matrix capture the cross-correlations between different bins of $\hat{\xi}_\pm$ one can look at the *correlation matrix*. This matrix is given in terms of the covariance matrix elements Cov_{ij} as

$$\mathrm{Corr}_{ij} = \frac{\mathrm{Cov}_{ij}}{\sqrt{\mathrm{Cov}_{ii}\mathrm{Cov}_{jj}}}\,. \tag{A1}$$

In Fig. A1 we compare the correlation matrix obtained from the lognormal model to the correlation matrix obtained from averaging all 50 jackknife estimates of the covariance matrix that were presented in Section 5, i.e. using setup I from Table 1. We show here the galaxy–jackknife with 400 re-samplings which divides the survey into the smallest sub-regions – the picture looks very similar for the pair scheme and for other numbers of re-samplings. The lower-left corner of Fig. A1 shows the autocorrelations of $\hat{\xi}_+$ and the upper-right corner shows the autocorrelations of $\hat{\xi}_-$. The upper-left and lower-right corners show the cross-correlations between the two correlation functions. Furthermore, the lower-right half of the plot shows the correlations obtained from the lognormal model and the upper-left half shows the correlations obtained from the average jackknife covariance estimate. Each column and row of pixels represents one angular bin and the bins range from 1 arcmin to 150 arcmin, starting on the lower-left corner.

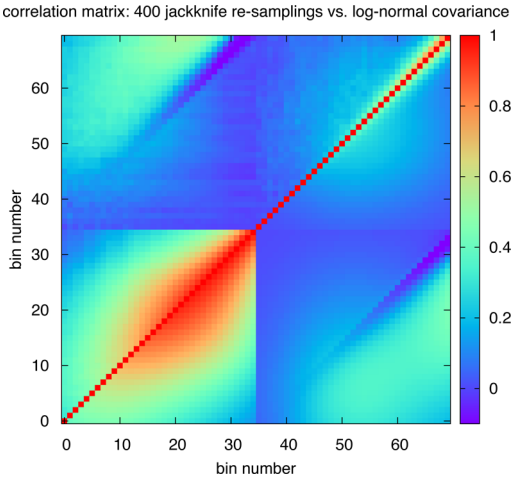

Figure A1. Correlation matrix of $\hat{\xi}_\pm$. Bins 0 to 34 belong to $\hat{\xi}_+$ and bins 35 to 69 belong to $\hat{\xi}_-$. For each two-point function the bins range from 1 arcmin to 150 arcmin, starting on the lower-left corner. The lower-right half of the plot displays the correlation coefficients of $\hat{\xi}_\pm$ obtained from the lognormal model and the upper-left halt displays the correlation coefficients obtained from the average jackknife covariance estimate (using 400 re-samplings).

Fig. A1 indicates that the jackknife estimator is able to capture the general structure of the correlation matrix of the two-point correlation functions. Given that internal covariance estimators mostly underestimate the variance of $\hat{\xi}_\pm$ one can hence conclude that the covariance elements Cov_{ij} are underestimated by the approximately same amount as the square root of $\mathrm{Cov}_{ii}\mathrm{Cov}_{jj}$ (cf. equation A1). This is, however, just a qualitative statement. We refer the reader to our comparison of parameter constraints in Section 5 for an application-related test of the performance of the jackknife estimators.

APPENDIX B: LIKELIHOOD CONTOURS

Figs B1 and B2 show the 1σ and 2σ contours in the Ω_m–σ_8 plane computed with COSMOLIKE when using galaxy–jackknife and pair–jackknife to estimate the covariance matrix (red and blue lines) and compare them to the same contours obtained from the true covariance matrix (black lines). The simulations are configured to mimic the complete 5-yr DES (cf. Section 5.2 or Table 1, setup IIb). The only thing that differs from simulation to simulation is the random seed that was used to generate the lognormal fields and the shape noise. The green dots represent the input cosmology of the simulations.

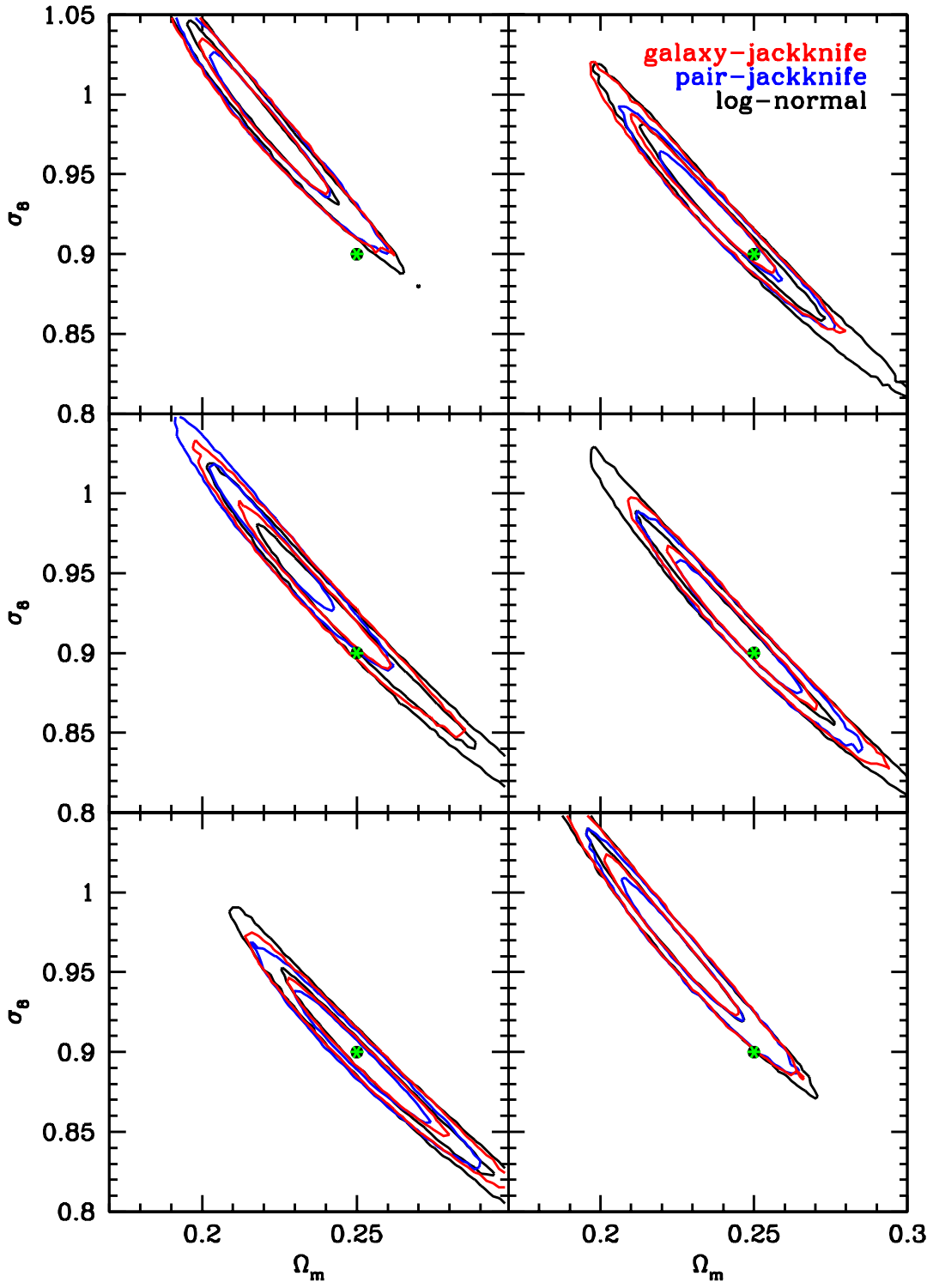

Figure B1. 1σ and 2σ contours in the $\Omega_{\rm m}$–σ_8 plane obtained from the first six simulations of setup IIb.

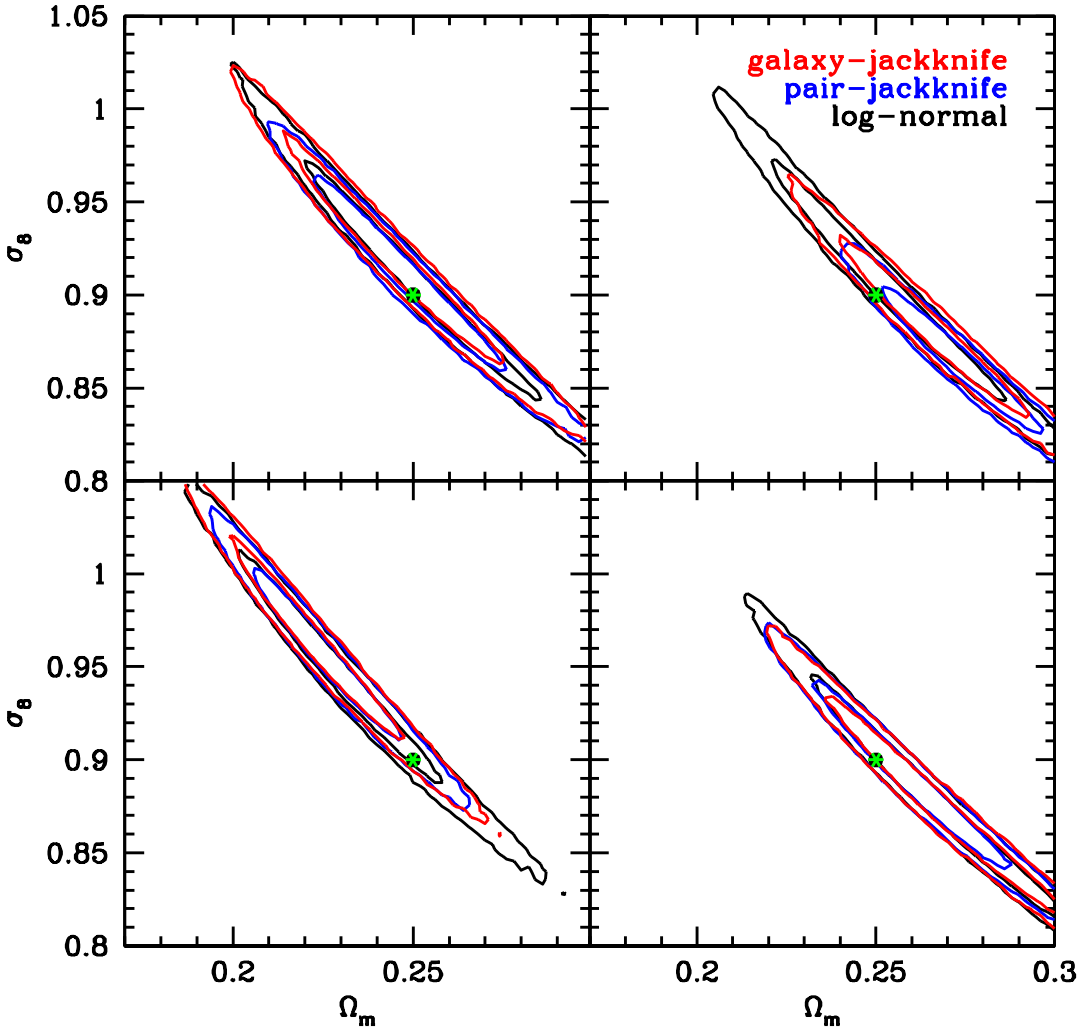

Figure B2. 1σ and 2σ contours in the Ω_m–σ_8 plane obtained from the remaining four simulations of setup IIb.

This paper has been typeset from a TEX/LATEX file prepared by the author.